

QY 237 EKVGRIDKNSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTSKNTSTS 296  
DB 241 EKVGRIDKNSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTSKNTSTS 300  
QY 297 RTHSEVGNAGAEVHAFEDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTAD 356  
DB 301 RTHSEVGNAGAEVHAFEDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTAD 360  
QY 357 TARLANIRYNTGTAPIYVNLPTTSLVLGNQTLATIKADENQISQILAPNNYPSKNL 416  
DB 361 TARLANIRYNTGTAPIYVNLPTTSLVLGNQTLATIKADENQISQILAPNNYPSKNL 420  
QY 417 APIALNAOKDASSTPTNNYQFLEKTKQLRLDQDVYGNIAATYAFNGRVRVDTCGN 476  
DB 421 APIALNAOKDASSTPTNNYQFLEKTKQLRLDQDVYGNIAATYAFNGRVRVDTCGN 460  
QY 477 MSEVLPOIQTETARIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAPFNGNEP 536  
DB 461 MSEVLPOIQTETARIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAPFNGNEP 520  
QY 537 NGNLOYQGGKIDTEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRF 596  
DB 521 NGNLOYQGGKIDTEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRF 580  
QY 597 HYDRNNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVI 656  
DB 581 HYDRNNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVI 640  
QY 657 NDRYDMLNIISSROQKGFIDFKYNDKPLIYISPNYKVNVAATKENTIIINSENGDT 716  
DB 641 NDRYDMLNIISSROQKGFIDFKYNDKPLIYISPNYKVNVAATKENTIIINSENGDT 700  
QY 717 STNGIKKILIFSKGYEIG 735  
DB 701 STNGIKKILIFSKGYEIG 719

RESULT 9  
US-09-273-839A-8  
; Sequence 8, Application US/09273839A  
; Patent No. 6329156  
; GENERAL INFORMATION:  
; APPLICANT: Cirino, Nick M  
; APPLICANT: Jackson, Paul J  
; APPLICANT: Lehnert, Bruce E  
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface  
; FILE REFERENCE: S-89,662  
; CURRENT APPLICATION NUMBER: US/09/273,839A  
; CURRENT FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 8  
; TYPE: PRT  
; LENGTH: 288  
; ORGANISM: Bacillus anthracis  
US-09-273-839A-8

Query Match 33.8%; Score 1275; DB 4; Length 288;  
Best Local Similarity 100.0%; Pred. No. 4.4e-84;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 486 ETETARIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAPFNGNEPNIQYQOK 545  
DB 29 ETETARIIFNGKOLNVERRIAANVPSDPLETTKPDWTLKEALKIAPFNGNEPNIQYQOK 88  
QY 546 DITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRFHYDRNNIAV 605  
DB 89 DITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDKRFHYDRNNIAV 148  
QY 606 GADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRYDMLNI 665  
DB 149 GADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRYDMLNI 208

QY 666 SSLROQKGFIDFKYNDKPLIYISPNYKVNVAATKENTIIINSENGDTSTNGIKKIL 725  
DB 209 SSLROQKGFIDFKYNDKPLIYISPNYKVNVAATKENTIIINSENGDTSTNGIKKIL 268  
QY 726 IFSKGYEIG 735  
DB 269 IFSKGYEIG 278

## RESULT 10

US-08-960-780-32  
; Sequence 32, Application US/08960780  
; Patent No. 6204435  
; GENERAL INFORMATION:  
; APPLICANT: Feitelson, Jerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schmeits, James  
; APPLICANT: Loewer, David  
; APPLICANT: Dullum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide  
; NUMBER OF SEQUENCES: 134  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606-6669  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/960,780  
; FILING DATE: 30-OCT-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,848  
; FILING DATE: 30-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: MA-708  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 881 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: 177C8  
US-08-960-780-32

Query Match 20.3%; Score 767; DB 3; Length 881;  
Best Local Similarity 30.1%; Pred. No. 8.1e-47;  
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;  
QY 4 QENRLLNESSSQGLGYFSDINFOAPMVVTSSTTGDLSIPSELEN--IPSENYFQ 61  
DB 39 QKNQ---QKMDRXGLLGYFKGKDF-SNLTMEAPTRDSTLIYDQQTANKLLDKKQOEYQ 94  
QY 62 SAISGFIKKYKSDYEYFATSDHVTWVDDQEVINKASNKIRLEKRLYQIKIQYQ 121

Db 95 SIRWGLIOSKETGDTFFNLSDDEQAIIEINGKISNKGKEQVHVHLEKGLVPKIEYQ 154  
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISDNLQLPKQKSS-----N 162  
Db 155 SD--TKFNIDSKTFKELKFKIDSONQPOQVQOQDELNPEFNKESQBFLLAKPSKINLFT 212  
QY 163 SRKQSTAGTPVPRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYSS 222  
Db 213 QMKKEIDEID--TDTGDSIPDLWEENGYTI---QNRIVKWDLSL-ASKGYTKFVSN 264  
QY 223 PEKWTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTONT 282  
Db 265 PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAAPSVNVSMKVLSPNENLS---- 320  
QY 283 DSETRISKNSTSTRTSTHTSEVHGNAEVAHAFDGGVSAGFSNSNS---TVAIDHLS 339  
Db 321 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 363  
QY 340 LAGERTWAETMG---LNTADTARLNANIRVYNTGTAPIYVNLPTTSLVGLKNOTLATIK 395  
Db 364 VAQE--WGTSTGNTSQFTASAGYLNANVRVNVGTGAIYDKVPTSPVL--NNDTIAIT 420  
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAQKASSTPTIMVNYQFLEKTKQLRDLTDQV 455  
Db 421 AKSNSTALNISPGESYKKGQNGIAITSMDDFNHSHPTITLNKKQVDNLLNKKPMLETNQT 480  
QY 456 YGNIAATYFNGRVYDVTGNSWSEVLPIQIETTARIIFNGKOLNLVERRIAAVNPSDPLE 515  
Db 481 DG---VYKIDKTHGNTVGGEMNGVIQIKAKTASIIVDGGE-RVAEKRAAKDYENBED 536  
QY 516 TTPDMLTKALKIAP--GENEFGNLQYQKIDTEFDF--NFDQOTSQNIKNQALAE-- 569  
Db 537 KT-PSLTKDALKSLSPDEIKETIEGLLYKNKPIYESVMTYLDNTAKEVTKQLNDTGT 595  
QY 570 ---NATNIYTVLKDILNAQWNLIRDKRPHYDRNNIAVCADESVVKEAHREVINSTREG 626  
Db 596 KFXDVSHLYDV---KLTPOKNTWIK-LSILYDN---ABSNDNSIGKWTNIVSGGNG 647  
QY 627 -----LLNID-----KOIRKILSYVIEIDE-----GLKE 654  
Db 648 KKQYSSNNPDANLTNTDAQKLNKRDYIISLYMKSEKNTQCEITRIDGEIYPIITKTVN 707  
QY 655 VINDRYDMLNT--SSLRQDCKTFIDPKYNDKPLVSIKPNYKVVYAVTKENTINPSE 712  
Db 708 VKNKYKRLDLIAHNKSNPISIIHKT-NDEITLFWDDLSI--TDVASIKPEN--LTDSE 763  
QY 713 NGDT-STNGIK---KLIFSKKGYEIG 735  
Db 764 IKQIYSRYGKLEDGILIDKKGTHYG 790

RESULT 11  
US-09-073-898-32  
; Sequence 32, Application US/09073898  
; Patent No. 6242669  
; GENERAL INFORMATION:  
; APPLICANT: Feitelson, Jerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schneits, James  
; APPLICANT: Loewer, David  
; APPLICANT: Dullum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; APPLICANT: Morrill, George  
; APPLICANT: Finstad-Lee, Stacy  
; TITLE OF INVENTION: No. 6242669a1 Pesticidal Toxins and Nucleotide  
; TITLE OF INVENTION: Sequences Which Encode These Toxins  
; NUMBER OF SEQUENCES: 144  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/073,898  
APPLICATION NUMBER: US/09/073,898  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,848  
FILING DATE: 30-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/960,780  
FILING DATE: 30-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-708C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 881 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: PS177C8  
US-09-073-898-32  
Query Match 20.3%; Score 767; DB 3; Length 881;  
Best Local Similarity 30.1%; Pred. No. 8.1e-47;  
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;  
QY 4 QENLLNESSSSQGLLYPFDLNFQAPMVVTSSTTGDLSIPSELEN--IPSENYQFQ 61  
Db 39 QKNQ--QKEMDRKGLLYYFKGKDF-SNLTMPAPTRDSTLIYDQQTANKLLDKKQOEY 94  
QY 62 SAISWGFYKVKSEYFATPSADNHVTWVDDQVINKASNSKIRLEKGLYQIKIYO 121  
Db 95 SIRWGLIOSKETGDTFFNLSDDEQAIIEINGKISNKGKEQVHVHLEKGLVPKIEY 154  
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISDNLQLPKQKSS-----N 162  
Db 155 SD--TKFNIDSKTFKELKFKIDSONQPOQVQOQDELNPEFNKESQBFLLAKPSKINLFT 212  
QY 163 SRKQSTAGTPVPRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYSS 222  
Db 213 QMKKEIDEID--TDTGDSIPDLWEENGYTI---QNRIVKWDLSL-ASKGYTKFVSN 264  
QY 223 PEKWTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTONT 282  
Db 265 PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAAPSVNVSMKVLSPNENLS---- 320  
QY 283 DSETRISKNSTSTRTSTHTSEVHGNAEVAHAFDGGVSAGFSNSNS---TVAIDHLS 339  
Db 321 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 363  
QY 340 LAGERTWAETMG---LNTADTARLNANIRVYNTGTAPIYVNLPTTSLVGLKNOTLATIK 395  
Db 364 VAQE--WGTSTGNTSQFTASAGYLNANVRVNVGTGAIYDKVPTSPVL--NNDTIAIT 420  
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAQKASSTPTIMVNYQFLEKTKQLRDLTDQV 455  
Db 421 AKSNSTALNISPGESYKKGQNGIAITSMDDFNHSHPTITLNKKQVDNLLNKKPMLETNQT 480



QY 456 YGNATYFNGRVRVDTGNSWSEVLPOIQETTARIIFNGKDLNVLVERIAAVNPSDPLE 515  
Db 481 DG---VYKIDTHGNIVTGEWNGVIOQIRAKTASIIIVDDGE-RVAEKRVAAKDYENPED 536  
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLQYQKDIETEDF--NFDQOTSONIKNQLAEL-- 569  
Db 537 KT-PSLTLDKALKSYDPEIETEGLLYKKNKFIYESSVMTYLDENTAKEVTKQLNDTTG 595  
QY 570 ---NATNIYTVLTKILNKAOWNLIRDRKPHYDRNNIAGVADSVVKEAHRVINSSTEG 626  
Db 596 KFKDVSHLYDV---KLTTPKMNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647  
QY 627 -----LLLNID-----KDIRKILSGYVIEIDTE-----GLKE 654  
Db 648 KKQYSSNNPDANLTNTDAQEKLNKNRDYIISLYMKSEKNTQCEITIDGEIYPTTKTVN 707  
QY 655 VINDRYDMLNI--SSLRQDGKTFIDFKYNDKULPLYISNPYKNVYAVTKENTINPSE 712  
Db 708 VNKDNYKRLDIIAHNIKSNPISSIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSE 763  
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735  
Db 764 IKQIYRYGKLEDDGILDKKGIHYG 790

RESULT 12  
US-09-307-106-8  
; Sequence 8, Application US/09307106  
; Patent No. 6603063  
; GENERAL INFORMATION:  
; APPLICANT: Fetteison, Jerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schmeits, James  
; APPLICANT: Loewer, David  
; APPLICANT: Dullum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; APPLICANT: Finstad-Lee, George  
; TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606-6669  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/307,106  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,848  
; FILING DATE: 30-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/960,780  
; FILING DATE: 30-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/073,898  
; FILING DATE: 05-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sanders, Jay M.  
; REGISTRATION NUMBER: 39,355

; REFERENCE/DOCKET NUMBER: MA-708C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 881 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: PS177C8a  
US-09-307-106-8

Query Match 20.3%; Score 767; DB 4; Length 881;  
Best Local Similarity 30.1%; Pred. No. 8.1e-47;  
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;

QY 4 QENLLNESSSOGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSSSELEN--IPSENYFQ 61  
Db 39 QKNQ---QKEMDRKGLLYGYFKGKDF-SNLTMPAPTRDSTLIYDQQTANKLLDKKQCEYQ 94  
QY 62 SAIWSGFIKVKSDSEYTFATSDAHNVMTWDDDEVINKASNSNKIRLEKRLYQIKIYQ 121  
Db 95 SIRWIGLIQSKETGDTFNLSDEQAIIIEINGKIIISNKGEKQVHVLEKGLVPIKIEYQ 154  
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISSNLOLPELKOKSS-----N 162  
Db 155 SD--TKFNIDSKTFKELKLFKIDSQNPQVQODELENPEFNKESQEFIAKPSKINLFT 212  
QY 163 SRKRTSAGTVPDRDNDGIPDSLEVEGYTVDPYKNTFLSPWISNIHKKGLTKYKSS 222  
Db 213 QMKREIDED---TDTGDSIPDLWEENGVTI-----QNRIAVKWDDSL-ASKGYTKFVSN 264  
QY 223 PEKWTASDPYSDEKVTGRIDKNVSPEARHPLVAAYPVHVDMENIILSKNEDOSTQNT 282  
Db 265 PLESHTVGDPTYDYEKAARDLDSNAKETENPLVAAPSVNVSMKEVILSPNENLS---- 320  
QY 283 DSETRTISKNTSTSTRTHTSEVHGNAEYHASFFDIGGSVAGFSNSNSS---TVAIDHSL 339  
Db 321 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVMYQHS 363  
QY 340 LAGERTWAETMG-----LNTADTARLNANIRVMTGTAPIYVNLPTSLVLGKQQTATIK 395  
Db 364 VAQE--WGTSSTGTSQNTASAGYLNANVRNNVGTGAIYDVKPTTSFVL--NNDTIATIT 420  
QY 396 ADENQLSQILAPNNYPSKNLAPIALNAQKADASTPITMNYNQFLEKTKQLRLDQV 455  
Db 421 AKSNSTALNTSPGSSYPKQGNGIAITSMDDFNSHPITLNNKQVDNLLNNKPMMLNQT 480  
QY 456 YGNATYFNGRVRVDTGNSWSEVLPOIQETTARIIFNGKDLNVLVERIAAVNPSDPLE 515  
Db 481 DG---VYKIDTHGNIVTGEWNGVIOQIRAKTASIIIVDDGE-RVAEKRVAAKDYENPED 536  
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLQYQKDIETEDF--NFDQOTSONIKNQLAEL-- 569  
Db 537 KT-PSLTLDKALKSYDPEIETEGLLYKKNKFIYESSVMTYLDENTAKEVTKQLNDTTG 595  
QY 570 ---NATNIYTVLTKILNKAOWNLIRDRKPHYDRNNIAGVADSVVKEAHRVINSSTEG 626  
Db 596 KFKDVSHLYDV---KLTTPKMNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647  
QY 627 -----LLLNID-----KDIRKILSGYVIEIDTE-----GLKE 654  
Db 648 KKQYSSNNPDANLTNTDAQEKLNKNRDYIISLYMKSEKNTQCEITIDGEIYPTTKTVN 707  
QY 655 VINDRYDMLNI--SSLRQDGKTFIDFKYNDKULPLYISNPYKNVYAVTKENTINPSE 712  
Db 708 VNKDNYKRLDIIAHNIKSNPISSIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSE 763  
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735

Db	764	IKQIVSRYGKLEDDGILLDKGGIHYG	790
RESULT 13			
US-09-850-351A-32			
; Sequence 32, Application US/09850351A			
; Patent No. 6656908			
GENERAL INFORMATION:			
APPLICANT: Fetteison, Jerald S.			
; Schnepf, H. Ernest			
; Narva, Kenneth E.			
; Stockhoff, Brian A.			
; Schneits, James			
; Loewer, David			
; Mullum, Charles Joseph			
; Muller-Cohn, Judy			
; Stamp, Lisa			
; Morrill, George			
TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide			
Sequences Which Encode These Toxins			
NUMBER OF SEQUENCES: 144			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik			
STREET: 2421 N.W. 41st Street, Suite A-1			
CITY: Gainesville			
STATE: FL			
COUNTRY: US			
ZIP: 32606-6669			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: Patent In Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/09/850,351A			
FILING DATE: 07-May-2001			
CLASSIFICATION: <Unknown>			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: US 09/073,898			
FILING DATE: 06-MAY-1998			
APPLICATION NUMBER: US 08/960,780			
FILING DATE: 30-OCT-1997			
APPLICATION NUMBER: US 60/029,848			
FILING DATE: 30-OCT-1996			
ATTORNEY/AGENT INFORMATION:			
NAME: Sanders, Jay M.			
REGISTRATION NUMBER: 39,355			
REFERENCE/DOCKET NUMBER: MA-708CD1			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 352-375-8100			
TELEFAX: 352-372-5800			
INFORMATION FOR SEQ ID NO: 32:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 881 amino acids			
TYPE: amino acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: peptide			
ORIGINAL SOURCE:			
INDIVIDUAL ISOLATE: PSI170C8			
SEQUENCE DESCRIPTION: SEQ ID NO: 32:			
US-09-850-351A-32			
Query Match 20.3%; Score 767; DB 4; Length 881;			
Best Local Similarity 30.1%; Pred. No. 8,1e-47;			
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;			
Qy	4	QENRLNESSSQGLLYGYSFDELNFPQAMVTSSTTGDLSPSSLEN--IPSENQYFQ	61
Db	39	QKNQ---QKEMDRKGLGYFFGKDP-SNLTWFAPTRSTLIYDQQTANKLDDKQBYQ	94
Qy	62	SAIWSGFIKVKSDYTFATSDNHNVTMWDDQEVINKASNNKRLKGRLYQIKIYQ	121



; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 884 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-471-044-5

```
Query Match      20.3%; Score 765; DB 2; Length 884;
Best Local Similarity 30.1%; Pred. No. 1.le-46;
Matches 245; Conservative 129; Mismatches 295; Indels 146; Gaps 34;

QY 4 QENRLNBSSESSQGLGYFSLNPAQPMVVTSTTGLSPSSSELEN--IPSENQYFQ 61
Db 42 QKNQ--QKEMDRKGLGYFFKGF-SNLTMAFTRDSTLIYDQQTANKLDRKQOEYQ 97
QY 62 SAIWSGFIVKVKDEYFATSDAHNVMTWVDQEVINKASNGNKIRLEKGRLYQIKIOYQ 121
Db 98 SIHWIGLJOSKETGDTFNLSEDEQAIIEINGKIISNGKEKEQVHVHLEKGLVPIKIEYQ 157
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISEDNLQLPELKQKSS-----N 162
Db 158 SD--TKFNIDSKTFKELKLFKIDSQNPQVQOQDELNPFNKESQEFLAKPSKINLFT 215
QY 163 SRKRSSTAGTPVDRDNGIDPSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 216 QMKRREIDED---TDTGDSIPDLWEENGTYI---QNRIVKDDSL-ASKGYTKFVSN 267
QY 223 PEKWTASDPYPSEKVTGRIDKNSPEARHPLVAAPIVHVDMENIILSKNEDOSTQNT 282
Db 268 PLESHTVGPDPYDYKAAARDLDSNAKETNPLVAAFPSPVSVMEKVLSPNENLS--- 323
QY 283 DSETRISKNTSRTHTSETHVGHAEVHASFFDIGGSVAGSAGFSNSNS---TVAIDHLS 339
Db 324 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 366
QY 340 LAGERTWAETMG---LNTACTARLNANRYVNTGAPYINVLPTTSLVLGKNQTLATIK 395
Db 367 VAQE--WGTSTGNTSQFNTASAGYLNANRYVNVGTGAIYDVKFTTSFVL--NNDTIATIT 423
QY 396 ADENQLSQILAPNNYFESKNLAPIALNAQKDASTPTWNYNQFLEKTKQLRLDTPQV 455
Db 424 AKSNSTALNTPSGESVPKGGQNGIAITSMDDFNSHPITLNKKQVDNLNKKPMLETNQT 483
QY 456 YGNIATYFNGRVRVDTGSNWSBEVLPIQIETTARIIFNGKOLNLVERRIAAVNPSPDPLE 515
Db 484 DG---VYKIKDTHGNVTGGWNGVIQQLKAKTASIIIVDDGE-RVAEKVRAKDYENPED 539
QY 516 TTKPDMTLKALKIAP--GFNEENGMLQYCGKDIETEDF--NPDOOTSQNIKNQLAEL-- 569
Db 540 KT-PSLTLDKALKSLVPDEIKEIEGLLYKKNKDIYESVNTYLDENTAKEVTKQLNDITG 598
QY 570 ---NATNIYTVLDKIKLNAQNILIRDKRPHYDRNNIAVGADESVVVKEAHREVINSSTEG 626
Db 599 KFKDVSHLYDV---KLTPOWNTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGNG 650
QY 627 -----LLNID-----KQIRKILSGYIVEIEDTE-----GLKE 654
Db 651 KKOYSSNNPDANLTNTDAQEKLNQRDYIISLYMKSEKNTQCEITIDGIEIYPIITTKTVN 710
QY 655 VINDRYDMLN-----ISSLRQDGKTFIDFKYNDKPLPIYISNPVKYNNVAVTKE 704
Db 711 VNKNYKRLDIIAHNIKSNFISSUH-----IKTNDIEITLFWDDISI-TDVASIKPE 760
QY 705 NTIINPSENGDT-STNGIK---KILIFSCKGYEIG 735
Db 761 N--LTDSEIKQIVSRYGKLEGDGILIDKKGHIHG 793
```

Search completed: May 3, 2004, 19:42:47  
Job time : 17.0986 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 36.6393 Seconds  
(without alignments)  
5560.545 Million cell updates/sec

Title: US-09-848-909A-13

Perfect score: 3772  
Sequence: 1 EVKQENLLNESSSSQGLL.....TSTNGIKLILFSKKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/prodata/2/pubpaa/US09D\_NEW\_PUB.pep:\*  
13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/2/pubpaa/US10D\_NEW\_PUB.pep:\*  
17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3753	99.5	735	15	US-10-410-647-30
2	3753	99.5	736	12	US-09-848-909-1
3	3753	99.5	736	12	US-09-848-909-2
4	3753	99.5	736	12	US-09-848-909-3
5	3753	99.5	736	12	US-09-848-909-4
6	3753	99.5	736	12	US-09-848-909-5
7	3753	99.5	736	12	US-09-848-909-6
8	3753	99.5	736	12	US-09-848-909-7
9	3753	99.5	736	12	US-09-848-909-8
10	3753	99.5	736	12	US-09-848-909-9
11	3753	99.5	736	12	US-09-848-909-10
12	3753	99.5	736	12	US-09-848-909-11
13	3753	99.5	736	12	US-09-848-909-12
14	3753	99.5	736	12	US-09-848-909-13
15	3753	99.5	736	12	US-09-848-909-14

16	3753	99.5	736	12	US-09-848-909-15	Sequence 15, Appl
17	3753	99.5	736	12	US-09-848-909-16	Sequence 16, Appl
18	3753	99.5	736	12	US-09-848-909-17	Sequence 17, Appl
19	3753	99.5	736	12	US-09-848-909-18	Sequence 18, Appl
20	3753	99.5	736	12	US-09-848-909-19	Sequence 19, Appl
21	3753	99.5	736	12	US-09-848-909-20	Sequence 20, Appl
22	3753	99.5	736	12	US-09-848-909-21	Sequence 21, Appl
23	3753	99.5	736	12	US-09-848-909-22	Sequence 22, Appl
24	3753	99.5	736	12	US-09-848-909-23	Sequence 23, Appl
25	3753	99.5	736	15	US-10-442-502-7	Sequence 7, Appl
26	3753	99.5	763	15	US-10-442-502-5	Sequence 5, Appl
27	3753	99.5	764	15	US-10-442-502-6	Sequence 6, Appl
28	3750	99.4	735	12	US-10-402-466A-9	Sequence 9, Appl
29	3750	99.4	735	12	US-10-402-466A-13	Sequence 13, Appl
30	3746	99.3	735	14	US-10-332-282-13	Sequence 13, Appl
31	3746	99.3	735	12	US-09-848-909-30	Sequence 30, Appl
32	3746	99.3	764	12	US-10-253-286-681	Sequence 681, App
33	3646	96.7	764	15	US-10-245-871-681	Sequence 681, App
34	3646	96.7	764	9	US-09-747-521-4	Sequence 4, Appl
35	3646	96.7	764	13	US-10-106-014-4	Sequence 4, Appl
36	3646	96.7	764	13	US-10-105-695-4	Sequence 4, Appl
37	3042	80.6	599	12	US-10-105-694-4	Sequence 24, Appl
38	3037	80.5	595	14	US-10-332-282-11	Sequence 11, Appl
39	2910	77.1	573	12	US-10-402-466A-22	Sequence 22, Appl
40	2898	76.8	569	15	US-10-442-502-8	Sequence 8, Appl
41	2488	66.0	487	14	US-10-332-282-5	Sequence 5, Appl
42	2173	57.6	426	14	US-10-332-282-9	Sequence 9, Appl
43	2127	56.4	423	12	US-10-402-466A-24	Sequence 24, Appl
44	1624	43.1	318	14	US-10-332-282-7	Sequence 7, Appl
45	1338	35.5	258	14	US-10-332-282-3	Sequence 3, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-410-647-30  
; Sequence 30, Application US/10410647  
; Publication No. US20030235818A1  
; GENERAL INFORMATION:  
; APPLICANT: PLEXUS VACCINE, INC.  
; APPLICANT: Katritch, Vsevolod  
; APPLICANT: Bordner, Andrew  
; APPLICANT: Deans, Robert  
; APPLICANT: Sumner, Mary  
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME  
; FILE REFERENCE: PLEX1110-1  
; CURRENT APPLICATION NUMBER: US/10/410,647  
; CURRENT FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: US 60/373,668  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: US 60/371,256  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: US 60/371,250  
; PRIOR FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 30  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-10-410-647-30

Query Match 99.5%; Score 3753; DB 15; Length 735;  
Best Local Similarity 99.6%; Pred. No. 2.8e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSSSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
DB 1 EVKQENLLNESSSSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
QY 61 QSAIWSGFVKVKSDEYTFATSDNHTVMVDQEVNKNASNKIKLEGRLYQIKIY 120

Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIQY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKXSSNRKXKSTASGPTVPDRN 180  
Db 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKXSSNRKXKSTASGPTVPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
QY 361 NANIRYNTGTAPTYNVLPPTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420  
Db 361 NANIRYNTGTAPTYNVLPPTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSKNWEV 480  
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSKNWEV 480  
QY 481 LPQOETARIIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPQOETARIIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNLIIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNLIIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPNKXNVAATKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPNKXNVAATKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

## RESULT 2

US-09-848-909-1  
; Sequence 1, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-1

Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. NO. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENLLNESBSSOGLLYFSDLNFOAPMVVTSITGDLSPSELENIPSENQYF 60  
Db 1 EVKQENLLNESBSSOGLLYFSDLNFOAPMVVTSITGDLSPSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIQY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIQY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKXSSNRKXKSTASGPTVPDRN 180  
Db 121 QRENPTKEGLDFKLYWTDSONKKEV:SSDNLQLOPELQKXSSNRKXKSTASGPTVPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTSETRTISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
QY 361 NANIRYNTGTAPTYNVLPPTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420  
Db 361 NANIRYNTGTAPTYNVLPPTSLVLGKNQTLATIKADENOLSQLAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSKNWEV 480  
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSKNWEV 480  
QY 481 LPQOETARIIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPQOETARIIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNLIIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNKXNLIIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPNKXNVAATKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPNKXNVAATKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

## RESULT 3

US-09-848-909-2  
; Sequence 2, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis

```
US-09-848-909-2
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-09-848-909-3

Query Match          99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKKRSTSGPTVDPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKKRSTSGPTVDPDRN 180
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPYDFEYKVT 240
DB 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWTASDPYDFEYKVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
DB 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDASSPTIMVNYNOFLEKTKQLRLDQVYGNATYNFENGRVRVDTGSNWSEV 480
DB 421 LNAQKDASSPTIMVNYNOFLEKTKQLRLDQVYGNATYNFENGRVRVDTGSNWSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
DB 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQKDIITEPFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKNWILIRKRFHYDR 600
DB 541 QYQKDIITEPFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKNWILIRKRFHYDR 600
QY 601 NNIAVGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIDTEGLKEVINDRY 660
QY 661 DMLNLSLRQDGKTFIDFKKYNDKPLIYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNLSLRQDGKTFIDFKKYNDKPLIYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4
US-09-848-909-3
; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04

US-09-848-909-4
; Sequence 4, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
```

; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastsEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-4

Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYGFFSDFLNFAQPMVVTSTTTGDLSPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNSESQGLLYGFFSDFLNFAQPMVVTSTTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIVKVKSDSYTFATSNADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIVKVKSDSYTFATSNADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRSTASDPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRSTASDPTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWTASDPTVPDRDN 240  
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWTASDPTVPDRDN 240  
QY 241 GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEQDQSTQNTDSETRTISKNTSSTRTHT 300  
DB 241 GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEQDQSTQNTDSETRTISKNTSSTRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYNNVLTPTSLVLGNQTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
DB 361 NANIRYNTGTAPIYNNVLTPTSLVLGNQTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
QY 421 LNAQKQDASSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480  
DB 421 LNAQKQDASSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENGNL 540  
DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENGNL 540  
QY 541 QYQKQDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600  
DB 541 QYQKQDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600  
QY 601 NNIAVGADESIVKAEHREVNSTEGILLNIDKIDKILSGYIVIEIDTGLKEVINDRY 660  
DB 601 NNIAVGADESIVKAEHREVNSTEGILLNIDKIDKILSGYIVIEIDTGLKEVINDRY 660  
QY 661 DMLNLSLRQDGTFFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720  
DB 661 DMLNLSLRQDGTFFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 6  
US-09-848-909-5

; Sequence 5, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastsEQ for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-5

Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYGFFSDFLNFAQPMVVTSTTTGDLSPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNSESQGLLYGFFSDFLNFAQPMVVTSTTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIVKVKSDSYTFATSNADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIVKVKSDSYTFATSNADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRSTASDPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRSTASDPTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWTASDPTVPDRDN 240  
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWTASDPTVPDRDN 240  
QY 241 GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEQDQSTQNTDSETRTISKNTSSTRTHT 300  
DB 241 GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEQDQSTQNTDSETRTISKNTSSTRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYNNVLTPTSLVLGNQTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
DB 361 NANIRYNTGTAPIYNNVLTPTSLVLGNQTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
QY 421 LNAQKQDASSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480  
DB 421 LNAQKQDASSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENGNL 540  
DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENGNL 540  
QY 541 QYQKQDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600  
DB 541 QYQKQDITEFDNFDDQSTQNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600  
QY 601 NNIAVGADESIVKAEHREVNSTEGILLNIDKIDKILSGYIVIEIDTGLKEVINDRY 660  
DB 601 NNIAVGADESIVKAEHREVNSTEGILLNIDKIDKILSGYIVIEIDTGLKEVINDRY 660  
QY 661 DMLNLSLRQDGTFFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720  
DB 661 DMLNLSLRQDGTFFIDFKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735



Db 721 IKKILIFSKGYEIG 735  
RESULT 7  
US-09-848-909-6  
; Sequence 6, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-6  
Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIVKKSDEVTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIVKKSDEVTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRSTASGTPVDPDRN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRSTASGTPVDPDRN 180  
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240  
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYNNVLPITSLVGLKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYNTGTAPIYNNVLPITSLVGLKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSNVEV 480  
Db 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSNVEV 480  
QY 541 QYQKDI TEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAKKNILIRDRKPHYDR 600  
Db 541 QYQKDI TEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAKKNILIRDRKPHYDR 600  
QY 601 NNTAVGADES VVKEAHEVINSSTEGLLNIDKIDKILSGYIVEEDTEGLKEVINDRY 660  
Db 601 NNTAVGADES VVKEAHEVINSSTEGLLNIDKIDKILSGYIVEEDTEGLKEVINDRY 660

QY 561 DMLNIISSLRQDGKTFIDFKYNDKLPYISNPYKYNVAVTAKENTIIINSENGDISTNG 720  
Db 561 DMLNIISSLRQDGKTFIDFKYNDKLPYISNPYKYNVAVTAKENTIIINSENGDISTNG 720  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735  
RESULT 8  
US-09-848-909-7  
; Sequence 7, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-7  
Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIVKKSDEVTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIVKKSDEVTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRSTASGTPVDPDRN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNSRKRSTASGTPVDPDRN 180  
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240  
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYNNVLPITSLVGLKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYNTGTAPIYNNVLPITSLVGLKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSNVEV 480  
Db 421 LNAQKDASSPTITWNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSNVEV 480  
QY 481 LPOIQUETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMTLKEALKIAPFGFNEPKNL 540  
Db 481 LPOIQUETARIIFNGKOLNLVERRIAANVPSDPLETTKPDMTLKEALKIAPFGFNEPKNL 540  
QY 541 QYQKDI TEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKNAKKNILIRDRKPHYDR 600

Db 541 QYQCKDITEFDNFDOQTSONIKNQLAELNATNIYTVLKDINKLNACNLIIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Db 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
QY 721 IKKILIFSFKKGYEIG 735  
Db 721 IKKILIFSFKKGYEIG 735  
RESULT 9  
US-09-848-909-8  
; Sequence 8, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-8  
Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQJPELKQKSSNRKGRSTAGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQJPELKQKSSNRKGRSTAGPTVPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAAPYIVHVDMENIILSKNEDQSTONTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAAPYIVHVDMENIILSKNEDQSTONTDSETRTISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 421 LNAQDASSPTITMNYNOFLEKTKQLRDTDOVGNATATYFNGRVRVDTGSNWSEV 480  
Db 421 LNAQDASSPTITMNYNOFLEKTKQLRDTDOVGNATATYFNGRVRVDTGSNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENGNL 540  
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIAGFNEPENGNL 540  
QY 541 QYQCKDITEFDNFDOQTSONIKNQLAELNATNIYTVLKDINKLNACNLIIRDKRPHYDR 600  
Db 541 QYQCKDITEFDNFDOQTSONIKNQLAELNATNIYTVLKDINKLNACNLIIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Db 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
QY 721 IKKILIFSFKKGYEIG 735  
Db 721 IKKILIFSFKKGYEIG 735  
RESULT 10  
US-09-848-909-9  
; Sequence 9, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-9  
Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQJPELKQKSSNRKGRSTAGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVITSSDNLQJPELKQKSSNRKGRSTAGPTVPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAAPYIVHVDMENIILSKNEDQSTONTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAAPYIVHVDMENIILSKNEDQSTONTDSETRTISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

```
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDASSTPTTMYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGNNWSEV 480
Db 421 LNAQDDFSTPTTMYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGNNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540
Qy 541 QYQGKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLAKMNLIRDRKPHYDR 600
Db 541 QYQGKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLAKMNLIRDRKPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODGTTFDFKYNKDKLPLYISNPNYKVNVAATYVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGTTFDFKYNKDKLPLYISNPNYKVNVAATYVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 11
US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-10

Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
```

```
Qy 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDASSTPTTMYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGNNWSEV 480
Db 421 LNAQDDFSTPTTMYNQFLELEKTKQLRLDQVYGNIAATYVNGRVRVDTGNNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540
Qy 541 QYQGKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLAKMNLIRDRKPHYDR 600
Db 541 QYQGKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLAKMNLIRDRKPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLINIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODGTTFDFKYNKDKLPLYISNPNYKVNVAATYVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGTTFDFKYNKDKLPLYISNPNYKVNVAATYVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735

RESULT 12
US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-11

Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2.9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
```

Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWTASDPYSDFEKT 240  
Qy 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRITSKNTSRHT 300  
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRITSKNTSRHT 300  
Qy 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLTADTARL 360  
Db 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLTADTARL 360  
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENOLSQLAPNNYPSKQLAPIA 420  
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENOLSQLAPNNYPSKQLAPIA 420  
Qy 421 LNAQKDASSPTITWYNOFLFLEKTKQLRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480  
Db 421 LNAQKDASSPTITWYNOFLFLEKTKQLRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
Qy 541 QYQKDIETEDFNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Db 541 QYQKDIETEDFNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Qy 601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRODGTFFIDFKYNDKPLVYISNPNYKNNVAVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGTFFIDFKYNDKPLVYISNPNYKNNVAVTKNTIINPSENGDTSTNG 720  
Qy 721 IKKILIPSKKGYEIG 735  
Db 721 IKKILIPSKKGYEIG 735

RESULT 13  
US-09-848-909-12  
; Sequence 12, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; TYPE: PRT  
; LENGTH: 736  
; ORGANISM: Bacillus anthracis  
US-09-848-909-12

Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
Qy 61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWYDQEVINKASNNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSDNHNVTMWYDQEVINKASNNKIRLEKGRLYQIKIY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKKRSTTSAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKKRSTTSAGTVPDRDN 180  
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWTASDPYSDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWTASDPYSDFEKT 240  
Qy 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRITSKNTSRHT 300  
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEQSTQNTDSETRITSKNTSRHT 300  
Qy 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLTADTARL 360  
Db 301 SEVHGNAEVAHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLTADTARL 360  
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENOLSQLAPNNYPSKQLAPIA 420  
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENOLSQLAPNNYPSKQLAPIA 420  
Qy 421 LNAQKDASSPTITWYNOFLFLEKTKQLRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480  
Db 421 LNAQKDASSPTITWYNOFLFLEKTKQLRLDTPQVYGNATYNFENGRVRVDTGSNWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
Qy 541 QYQKDIETEDFNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Db 541 QYQKDIETEDFNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Qy 601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRODGTFFIDFKYNDKPLVYISNPNYKNNVAVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGTFFIDFKYNDKPLVYISNPNYKNNVAVTKNTIINPSENGDTSTNG 720  
Qy 721 IKKILIPSKKGYEIG 735  
Db 721 IKKILIPSKKGYEIG 735

RESULT 14  
US-09-848-909-13  
; Sequence 13, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; TYPE: PRT  
; LENGTH: 736  
; ORGANISM: Bacillus anthracis  
US-09-848-909-13

Query Match 99.5%; Score 3753; DB 12; Length 736;  
Best Local Similarity 99.6%; Pred. No. 2.9e-274;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60

```
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFTKVKKSDDEYTFATADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
Db 61 QSAIWSGFTKVKKSDDEYTFATADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGTTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGTTPVDRDN 180
Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASTASDPYDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Qy 301 SEVHGNAEVAHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLTADTARL 360
Db 301 SEVHGNAEVAHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYYPSKNLAPIA 420
Qy 421 LNAQKDSASPTIMTNYNQFLEKTKQLRLDQVYGNATYTNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSASPTIMTNYNQFLEKTKQLRLDQVYGNATYTNFENGVRVDTGSNWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRY 660
Db 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication NC. US2002003958A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/050002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-14
```

```
Query Match 99.5%; Score 3753; DB 12; Length 736;
Best Local Similarity 99.6%; Pred. No. 2,9e-274;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGLSPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFTKVKKSDDEYTFATADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
Db 61 QSAIWSGFTKVKKSDDEYTFATADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGTTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQKSSNSRKRSTSGTTPVDRDN 180
Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASTASDPYDFEKT 240
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Qy 301 SEVHGNAEVAHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLTADTARL 360
Db 301 SEVHGNAEVAHAFDIDGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTMGLTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYYPSKNLAPIA 420
Qy 421 LNAQKDSASPTIMTNYNQFLEKTKQLRLDQVYGNATYTNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSASPTIMTNYNQFLEKTKQLRLDQVYGNATYTNFENGVRVDTGSNWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRY 660
Db 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

```
Search completed: May 3, 2004, 20:02:13
Job time : 38.6393 secs
```

1	3772	100.0	735	1	PCT-US03-35733-13	Sequence 13, Appl
2	3772	100.0	735	23	US-09-848-909A-13	Sequence 13, Appl
3	3766	99.8	735	1	PCT-US03-35733-10	Sequence 10, Appl
4	3766	99.8	735	23	US-09-848-909A-10	Sequence 10, Appl
5	3763	99.8	735	1	PCT-US03-19786-4	Sequence 4, Appl
6	3760	99.7	735	1	PCT-US03-35733-2	Sequence 2, Appl
7	3760	99.7	735	23	US-09-848-909A-2	Sequence 2, Appl
8	3759	99.7	735	1	PCT-US03-35733-8	Sequence 8, Appl
9	3759	99.7	735	1	PCT-US03-35733-9	Sequence 9, Appl
10	3759	99.7	735	23	US-09-848-909A-8	Sequence 8, Appl
11	3759	99.7	735	23	US-09-848-909A-9	Sequence 9, Appl
12	3755	99.5	735	1	PCT-US03-35733-7	Sequence 7, Appl
13	3755	99.5	735	1	PCT-US03-35733-23	Sequence 23, Appl
14	3755	99.5	735	23	US-09-848-909A-7	Sequence 7, Appl
15	3755	99.5	735	23	US-09-848-909A-23	Sequence 23, Appl
16	3754	99.5	735	1	PCT-US03-35733-4	Sequence 4, Appl
17	3754	99.5	735	1	PCT-US03-35733-6	Sequence 6, Appl
18	3754	99.5	735	1	PCT-US03-35733-18	Sequence 18, Appl
19	3754	99.5	735	23	US-09-848-909A-4	Sequence 4, Appl
20	3754	99.5	735	23	US-09-848-909A-6	Sequence 6, Appl
21	3754	99.5	735	23	US-09-848-909A-18	Sequence 18, Appl
22	3753	99.5	735	1	PCT-US03-35733-5	Sequence 5, Appl
23	3753	99.5	735	1	PCT-US03-35733-11	Sequence 11, Appl
24	3753	99.5	735	1	PCT-US03-35733-17	Sequence 17, Appl
25	3753	99.5	735	1	PCT-US03-35733-19	Sequence 19, Appl
26	3753	99.5	735	1	PCT-US03-35733-20	Sequence 20, Appl
27	3753	99.5	735	1	PCT-US03-35733-21	Sequence 21, Appl
28	3753	99.5	735	22	US-09-791-537-43735	Sequence 43735, A
29	3753	99.5	735	23	US-09-848-909A-5	Sequence 5, Appl
30	3753	99.5	735	23	US-09-848-909A-11	Sequence 11, Appl
31	3753	99.5	735	23	US-09-848-909A-17	Sequence 17, Appl
32	3753	99.5	735	23	US-09-848-909A-19	Sequence 19, Appl
33	3753	99.5	735	23	US-09-848-909A-20	Sequence 20, Appl
34	3753	99.5	735	23	US-09-848-909A-21	Sequence 21, Appl
35	3753	99.5	735	30	US-10-410-647-30	Sequence 30, Appl
36	3753	99.5	736	1	PCT-US01-14372A-1	Sequence 1, Appl
37	3753	99.5	736	1	PCT-US01-14372A-2	Sequence 2, Appl
38	3753	99.5	736	1	PCT-US01-14372A-3	Sequence 3, Appl
39	3753	99.5	736	1	PCT-US01-14372A-4	Sequence 4, Appl
40	3753	99.5	736	1	PCT-US01-14372A-5	Sequence 5, Appl
41	3753	99.5	736	1	PCT-US01-14372A-6	Sequence 6, Appl
42	3753	99.5	736	1	PCT-US01-14372A-7	Sequence 7, Appl
43	3753	99.5	736	1	PCT-US01-14372A-8	Sequence 8, Appl
44	3753	99.5	736	1	PCT-US01-14372A-9	Sequence 9, Appl
45	3753	99.5	736	1	PCT-US01-14372A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
PCT-US03-35733-13  
; Sequence 13, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-13

Query Match 100.0%; Score 3772; DB 1; Length 735;  
Best Local Similarity 100.0%; Pred. No. 1.8e-298;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds  
(without alignments)  
4086.665 Million cell updates/sec  
Title: US-09-848-909A-13  
Perfect score: 3772  
Sequence: 1 EVKQENRLNSESSESSQGLL.....TSTNGIKKILIFSKKGVBIG 735  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 6019581 seqs, 976053577 residues  
Total number of hits satisfying chosen parameters: 6019581  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /cgn2\_6/prodata/2/paa/pctus COMB.pcp.\*
- 2: /cgn2\_6/prodata/2/paa/us06 COMB.pcp.\*
- 3: /cgn2\_6/prodata/2/paa/us07 COMB.pcp.\*
- 4: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 5: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 6: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 7: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 8: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 9: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 10: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 11: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 12: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 13: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 14: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 15: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 16: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 17: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 18: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 19: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 20: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 21: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 22: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 23: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 24: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 25: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 26: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 27: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 28: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 29: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 30: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 31: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 32: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*
- 33: /cgn2\_6/prodata/2/paa/us08 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

QY 1 EVKQENRLNSESSESSQGLLYGFFDLNFQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180  
Db 121 QRENPTKEGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKASSTPTIMVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 480  
Db 421 LNAQKASSTPTIMVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 480  
QY 481 LPOIQTETARIIIFNGKDLNVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPOIQTETARIIIFNGKDLNVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
QY 601 NNIAVGADESIVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESIVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 2

US-09-848-909A-13  
; Sequence 13, Application US/09848909A

; GENERAL INFORMATION:  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis

## US-09-848-909A-13

Query Match 100.0%; Score 3772; DB 23; Length 735;  
Best Local Similarity 100.0%; Pred. No. 1.8e-298;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGFFDLNFQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGFFDLNFQAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180  
Db 121 QRENPTKEGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKKRSTASAGTVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKASSTPTIMVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 480  
Db 421 LNAQKASSTPTIMVYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 480  
QY 481 LPOIQTETARIIIFNGKDLNVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPOIQTETARIIIFNGKDLNVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
QY 601 NNIAVGADESIVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESIVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 3

PCT-US03-35733-10  
; Sequence 10, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 10
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-10

Query Match          99.8%; Score 3766; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 5.5e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSESSQGLGYFSDLNFAQPMVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSESSQGLGYFSDLNFAQPMVTSSTTGDLSIPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQLPKQKSSNRKRSSTAGTTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQLPKQKSSNRKRSSTAGTTPVDRDN 180
Qy 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDSSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480
Qy 481 LPQIQTETARIIFNGKOLNLVERRIAIAVNPSPDLETTKPDMTLKEALKAFGNENGNL 540
Db 481 LPQIQTETARIIFNGKOLNLVERRIAIAVNPSPDLETTKPDMTLKEALKAFGNENGNL 540
Qy 541 QYQKDIETFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
Db 541 QYQKDIETFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
Qy 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

RESULT 4  
US-09-848-909A-10  
; Sequence 10, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; OF INFECTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A

```
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-10

Query Match          99.8%; Score 3766; DB 23; Length 735;  
Best Local Similarity 99.9%; Pred. No. 5.5e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSESSQGLGYFSDLNFAQPMVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSESSQGLGYFSDLNFAQPMVTSSTTGDLSIPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQLPKQKSSNRKRSSTAGTTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNLQLPKQKSSNRKRSSTAGTTPVDRDN 180
Qy 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDSSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSSTPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNWSEV 480
Qy 481 LPQIQTETARIIFNGKOLNLVERRIAIAVNPSPDLETTKPDMTLKEALKAFGNENGNL 540
Db 481 LPQIQTETARIIFNGKOLNLVERRIAIAVNPSPDLETTKPDMTLKEALKAFGNENGNL 540
Qy 541 QYQKDIETFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
Db 541 QYQKDIETFDNFDQQTSONIKNQLAELNATNIYTVLDKIKLNAKKNILIRDKRFHYDR 600
Qy 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRQDKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

RESULT 5  
PCT-US03-19786-4  
; Sequence 4, Application PC/TUS0319786  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.



; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof  
; FILE REFERENCE: PV59SPCT  
; CURRENT APPLICATION NUMBER: PCT/US03/19786  
; CURRENT FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Mature PA sequence including an E7B signal sequence  
PCT-US03-19786-4

Query Match 99.8%; Score 3763; DB 1; Length 735;  
Best Local Similarity 99.7%; Pred. No. 9.7e-298;  
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy	1	EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Db	1	EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Qy	61	QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Db	61	QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Qy	121	QRENPEKGLDFKLYWTDSONKKEVIVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
Db	121	QRENPEKGLDFKLYWTDSONKKEVIVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
Qy	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
Db	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
Qy	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTSKNTSTSRTH	300
Db	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTSKNTSTSRTH	300
Qy	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Qy	361	NANIRYVNTGTAPYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA	420
Db	361	NANIRYVNTGTAPYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA	420
Qy	421	LNAQKDA SSTPI TWNYNQFLEKTKQLRLD TDQVYGNIA TYNFENGVRVDTGSNWSEV	480
Db	421	LNAQKDA SSTPI TWNYNQFLEKTKQLRLD TDQVYGNIA TYNFENGVRVDTGSNWSEV	480
Qy	481	LPQIQTETARIIFNGKDLNVERRIA AAVNPSPDPLETTKPDMTLKEALKIAFGFNEP	540
Db	481	LPQIQTETARIIFNGKDLNVERRIA AAVNPSPDPLETTKPDMTLKEALKIAFGFNEP	540
Qy	541	QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKI KLNKNNILIRDKRPHYDR	600
Db	541	QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKI KLNKNNILIRDKRPHYDR	600
Qy	601	NNIYAVGADES VVKEAHREVINSSTEGLLNIDKDIRKILSGYI VEIEDTEGLKEVIN	660
Db	601	NNIYAVGADES VVKEAHREVINSSTEGLLNIDKDIRKILSGYI VEIEDTEGLKEVIN	660
Qy	661	DMLNIISSLRQDKTFIDFKYNDKLPYISNPYKVNYYAVTKNTIINPSENGDTSTNG	720
Db	661	DMLNIISSLRQDKTFIDFKYNDKLPYISNPYKVNYYAVTKNTIINPSENGDTSTNG	720
Qy	721	IKKILIFSKKGYEIG	735
Db	721	IKKILIFSKKGYEIG	735

RESULT 6  
PCT-US03-35733-2

; Sequence 2, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-2

Query Match 99.7%; Score 3760; DB 1; Length 735;  
Best Local Similarity 99.7%; Pred. No. 1.7e-297;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Db	1	EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
Qy	61	QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Db	61	QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIQY	120
Qy	121	QRENPEKGLDFKLYWTDSONKKEVIVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
Db	121	QRENPEKGLDFKLYWTDSONKKEVIVSSDNLQIPELKQKSSNRKKGSTAGPTVPDRDN	180
Qy	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
Db	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPVSDPEKVT	240
Qy	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTSKNTSTSRTH	300
Db	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETRTSKNTSTSRTH	300
Qy	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Qy	361	NANIRYVNTGTAPYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA	420
Db	361	NANIRYVNTGTAPYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA	420
Qy	421	LNAQKDA SSTPI TWNYNQFLEKTKQLRLD TDQVYGNIA TYNFENGVRVDTGSNWSEV	480
Db	421	LNAQKDA SSTPI TWNYNQFLEKTKQLRLD TDQVYGNIA TYNFENGVRVDTGSNWSEV	480
Qy	481	LPQIQTETARIIFNGKDLNVERRIA AAVNPSPDPLETTKPDMTLKEALKIAFGFNEP	540
Db	481	LPQIQTETARIIFNGKDLNVERRIA AAVNPSPDPLETTKPDMTLKEALKIAFGFNEP	540
Qy	541	QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKI KLNKNNILIRDKRPHYDR	600
Db	541	QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKI KLNKNNILIRDKRPHYDR	600
Qy	601	NNIYAVGADES VVKEAHREVINSSTEGLLNIDKDIRKILSGYI VEIEDTEGLKEVIN	660
Db	601	NNIYAVGADES VVKEAHREVINSSTEGLLNIDKDIRKILSGYI VEIEDTEGLKEVIN	660
Qy	661	DMLNIISSLRQDKTFIDFKYNDKLPYISNPYKVNYYAVTKNTIINPSENGDTSTNG	720
Db	661	DMLNIISSLRQDKTFIDFKYNDKLPYISNPYKVNYYAVTKNTIINPSENGDTSTNG	720
Qy	721	IKKILIFSKKGYEIG	735
Db	721	IKKILIFSKKGYEIG	735

```
RESULT 7
US-09-848-909A-2
; Sequence 2, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-2

Query Match          99.7%; Score 3760; DB 23; Length 735;
Best Local Similarity 99.7%; Pred. No. 1.7e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKSDSYTATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKSDSYTATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSIRHT 300
Db 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSIRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOFLATIKADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOFLATIKADENQLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNIAATYVNFENGRVVRVDTGSNWEV 480
Db 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNIAATYVNFENGRVVRVDTGSNWEV 480
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600
QY 601 NNIAVGADESIVKAEHREVSINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESIVKAEHREVSINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKFLYISNPNKYVNVAVTKENTIINPSENGDSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKFLYISNPNKYVNVAVTKENTIINPSENGDSTNG 720

RESULT 8
PCT-US03-35733-8
; Sequence 8, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-8

Query Match          99.7%; Score 3759; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.1e-297;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKSDSYTATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKSDSYTATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSSTAGTVPDRDN 180
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSIRHT 300
Db 241 GRIDKNVSPPEARPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSIRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOFLATIKADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKNOFLATIKADENQLSOILAPNNYPSKNLAPIA 420
QY 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNIAATYVNFENGRVVRVDTGSNWEV 480
Db 421 LNAQKDSSTPTITWYVNFQLEKTKQLRLDQVYGNIAATYVNFENGRVVRVDTGSNWEV 480
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600
QY 601 NNIAVGADESIVKAEHREVSINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESIVKAEHREVSINSTEGLLLNIDKDIRKILSGYVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKFLYISNPNKYVNVAVTKENTIINPSENGDSTNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKFLYISNPNKYVNVAVTKENTIINPSENGDSTNG 720
```

Db 601 NNIAVGADSVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNISLRQDGTFFIDFKKYNNDKPLIYISNPNYKNNVAVTKENTIINPSNGDTSTNG 720  
Db 661 DMLNISLRQDGTFFIDFKKYNNDKPLIYISNPNYKNNVAVTKENTIINPSNGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735  
RESULT 9  
PCT-US03-35733-9  
; Sequence 9, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-9  
Query Match 99.7%; Score 3759; DB 1; Length 735;  
Best Local Similarity 99.7%; Pred. No. 2.1e-297;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYGYPFDLNFQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGYPFDLNFQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDAHNVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDAHNVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPLKQKSNRKKRSTSGAGTPVDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPLKQKSNRKKRSTSGAGTPVDRDN 180  
QY 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYDFEKT 240  
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENOLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENOLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGSGNWSEV 480  
Db 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGSGNWSEV 480  
QY 481 LPOIETTARIIFNGKDNLVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540  
Db 481 LPOIETTARIIFNGKDNLVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540  
QY 541 QYQKCDITEFFDNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600

Db 541 QYQKCDITEFFDNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600  
QY 601 NNIAVGADSVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADSVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNISLRQDGTFFIDFKKYNNDKPLIYISNPNYKNNVAVTKENTIINPSNGDTSTNG 720  
Db 661 DMLNISLRQDGTFFIDFKKYNNDKPLIYISNPNYKNNVAVTKENTIINPSNGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735  
RESULT 10  
US-09-848-909A-8  
; Sequence 8, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-8  
Query Match 99.7%; Score 3759; DB 23; Length 735;  
Best Local Similarity 99.7%; Pred. No. 2.1e-297;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYGYPFDLNFQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGYPFDLNFQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDAHNVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDAHNVTMVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPLKQKSNRKKRSTSGAGTPVDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNQLQPLKQKSNRKKRSTSGAGTPVDRDN 180  
QY 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYDFEKT 240  
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENOLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENOLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGSGNWSEV 480  
Db 421 LNAQKDDASSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGSGNWSEV 480  
QY 481 LPOIETTARIIFNGKDNLVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540

Db 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMLKEALKIAPGNEPGL 540  
QY 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAMNIIIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAMNIIIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODQKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

RESULT 11

US-09-848-909A-9  
; Sequence 9, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-9

Query Match 99.7%; Score 3759; DB 23; Length 735;  
Best Local Similarity 99.7%; Pred. No. 2, 1e-297;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 180  
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEKT 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
QY 301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVNTGTAPIYVNLPTTSVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYVNLPTTSVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQDASSTPIITMYNQFLEKTKOLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480  
Db 421 LNAQDASSTPIITMYNQFLEKTKOLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480  
QY 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMLKEALKIAPGNEPGL 540  
Db 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMLKEALKIAPGNEPGL 540  
QY 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAMNIIIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFQDOTSONIKQLAELNATNIYVLDKIKLNAMNIIIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODQKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

RESULT 12

PCT-US03-35733-7  
; Sequence 7, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; PRIOR FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-7

Query Match 99.5%; Score 3755; DB 1; Length 735;  
Best Local Similarity 99.6%; Pred. No. 4, 4e-297;  
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTTGDLSPSSSELENIPSENOYF 180  
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEKT 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
QY 301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAFPDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NAMIRYVNTGTAPIYNNVLPPTSLVLGKNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NAMIRYVNTGTAPIYNNVLPPTSLVLGKNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKASSTPIITMNNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
DB 421 LNAQEDFSSTPIITMNNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 481 LPQIQETTAIIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540  
DB 481 LPQIQETTAIIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540  
QY 541 QYQKDIITBDFNFDQDQTSQNKQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600  
DB 541 QYQKDIITBDFNFDQDQTSQNKQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
DB 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
QY 661 DMLNISSLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720  
DB 661 DMLNISSLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735  
RESULT 13  
PCT-US03-35733-23  
; Sequence 23, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; PRIOR FILING DATE: 2003-10-10  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 23  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
; NAME/KEY: VARIANT  
; LOCATION: 427  
; OTHER INFORMATION: Xaa = any amino acid except Phe  
PCT-US03-35733-23  
Query Match 99.5%; Score 3755; DB 1; Length 735;  
Best Local Similarity 99.6%; Pred. No. 4.4e-297;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKRIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKRIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKCRSTAGTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKCRSTAGTVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYSPEKWKSTASDPSDFKVT 240  
DB 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYSPEKWKSTASDPSDFKVT 240

QY 241 GRIDKNVSPPEARHPLVAAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NAMIRYVNTGTAPIYNNVLPPTSLVLGKNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NAMIRYVNTGTAPIYNNVLPPTSLVLGKNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKASSTPIITMNNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
DB 421 LNAQDDXSSSTPIITMNNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 481 LPQIQETTAIIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540  
DB 481 LPQIQETTAIIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540  
QY 541 QYQKDIITBDFNFDQDQTSQNKQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600  
DB 541 QYQKDIITBDFNFDQDQTSQNKQLAELNATNIYTVLDKIKLNAKNILLIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
DB 601 NNIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
QY 661 DMLNISSLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720  
DB 661 DMLNISSLRODQGTFFDFKKNYDKLPLYISNPNYKVNVAVTKENTIINPSENGTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735  
RESULT 14  
US-09-848-909A-7  
; Sequence 7, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-7  
Query Match 99.5%; Score 3755; DB 23; Length 735;  
Best Local Similarity 99.6%; Pred. No. 4.4e-297;  
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNSESSESSQGLLYYFSDLNFOAPMVVTSSTTGLDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKRIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKRIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKCRSTAGTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQJPELKQKSSNRKCRSTAGTVPDRDN 180

QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWASTASDPYSDPEKVT 240
Db 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWASTASDPYSDPEKVT 240
QY 241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
QY 421 LNAQKDSASTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSASTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
QY 481 LPQIOETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPQIOETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
QY 541 QYQKDIETFDNFDPDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQKDIETFDNFDPDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPYKNNVYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPYKNNVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
RESULT 15
US-09-848-909A-23
; Sequence 23 Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 427
; OTHER INFORMATION: Xaa = any amino acid except Phe
US-09-848-909A-23
Query Match 99.5%; Score 3755; DB 23; Length 735;
Best Local Similarity 99.6%; Pred. No. 4.4e-297;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 EVKQENRLLNSESQGLLYGYFSDLNFQAPVWVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 1 EVKQENRLLNSESQGLLYGYFSDLNFQAPVWVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAINSGFIKVKSSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLLEKGLYQIKIY 120
Db 61 QSAINSGFIKVKSSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRKSTASAGTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRKSTASAGTVPDRN 180
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWASTASDPYSDPEKVT 240
Db 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSWASTASDPYSDPEKVT 240
QY 241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPVAAYPIVHVDMENIILSKNEDQSTONTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPPSKNLAPIA 420
QY 421 LNAQKDSASTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDSASTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
QY 481 LPQIOETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPQIOETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540
QY 541 QYQKDIETFDNFDPDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQKDIETFDNFDPDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPYKNNVYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPYKNNVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
Search completed: May 3, 2004, 19:57:44
Job time : 177.546 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 Seconds  
(without alignments)  
5403.004 Million cell updates/sec

Title: US-09-848-909A-13  
Perfect score: 3772  
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKLIPFKKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 segs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*\*  
1: pir1:\*\*  
2: pir2:\*\*  
3: pir3:\*\*  
4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3753	99.5	764	2 I39934	protective antigen
2	868	23.0	875	2 I40862	iota toxin compone
3	235.5	6.2	192	2 I39933	cryptic protein -
4	235.5	6.2	204	2 G59104	hypothetical prote
5	202	5.4	4688	2 F82885	hypothetical prote
6	190.5	5.1	2401	2 T28676	rhostry protein -
7	189.5	5.0	2523	2 B64635	toxin-like outer m
8	186.5	4.9	1125	2 E90598	membrane nuclease,
9	184.5	4.9	1639	2 S05603	major merozoite su
10	183	4.9	6713	2 B89921	hypothetical prote
11	181	4.8	1302	1 JC6009	surface-located me
12	181	4.8	2269	2 T28677	rhostry protein -
13	179.5	4.8	1072	2 A86827	hypothetical prote
14	179	4.7	1193	2 S68218	botulinum neurotox
15	179	4.7	1227	2 C97033	uncharacterized pr
16	178	4.7	1635	2 A10452	hemolysin [importe
17	177.5	4.7	2399	2 H71879	toxin-like outer m
18	177	4.7	4919	2 T31105	hypothetical prote
19	176	4.7	752	2 G50599	hypothetical prote
20	174.5	4.6	1355	2 T30822	lmp1 protein - Myc
21	173	4.6	1939	2 T18372	repeat organellar
22	173	4.6	4152	2 T31102	filamentous hemag
23	172	4.6	1837	2 T41023	probable nuclear p
24	170	4.5	1308	2 E71622	probable membrane
25	169.5	4.5	2340	2 B71704	cell surface antig
26	169	4.5	1631	1 SAZOK1	major merozoite su
27	169	4.5	1658	2 S55101	hypothetical prote
28	169	4.5	5005	2 F82884	hypothetical prote
29	168.5	4.5	821	2 S67087	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

I39934  
Protective antigen precursor - Bacillus anthracis plasmid  
C:Species: Bacillus anthracis  
C:Date: 19-Jul-1996 #sequence, revision 19-Jul-1996 #text\_change 01-Dec-2000  
C:Accession: I39934; S69160; F59104  
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.  
Gene 69, 287-300, 1988  
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr  
A:Reference number: I39933; MUID:89172073; PMID:3148491  
A:Accession: I39934  
A:Status: preliminary; translated from GE/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-764 <RES>  
A:Cross-references: GB:M22589; NID:G143280; PIDN:AAA2637.1; PID:G143282  
R:Priedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klimpel, K.R.; Birch, N.P.; Loh, Y.P.  
Arch. Biochem. Biophys. 316, 5-13, 1995  
A:Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SE  
A:Reference number: S69160; MUID:95142670; PMID:7840657  
A:Accession: S69160  
A:Molecule type: protein  
A:Residues: 197-202 <PRI>  
R:Okunaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle  
J. Bacteriol. 181, 6509-6515, 1999  
A:Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbor  
A:Reference number: A59091; MUID:99445483; PMID:10515943  
A:Accession: F59104  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-313, 'Q', 315-764 <OKI>  
A:Cross-references: GB:AF065404; NID:G4894216; PIDN:AAD32414.1; PID:G4894326  
A:Experimental source: strain Sterne  
A:Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid  
C:Genetics:  
A:Gene: pXOI-110  
A:Genome: plasmid  
C:Function:  
A:Description: three component exotoxin; protective antigen binds to receptors on the s  
y active components edema factor or lethal factor; the complex is internalized by recep  
C:Keywords: exotoxin  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-196/Domain: propeptide #status predicted <PRO>  
F:197-202/Product: protective antigen #status experimental <MAT>

Query Match 99.5%; Score 3753; DB 2; Length 764;  
Best Local Similarity 99.6%; Pred. No. 1.5e-182;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 EVKQENRLNSESQGLLYFSDLNFOAPMVVTSITGDLSPSELENIPSENYQF 60  
|||||  
Db 30 EVKQENRLNSESQGLLYFSDLNFOAPMVVTSITGDLSPSELENIPSENYQF 89

probable coiled-co  
alpha-toxin - Clos  
surface membrane p  
trse-like protein  
TyB protein - yeas  
botulinum neurotox  
hypothetical prote  
TyB protein - yeas  
TyB protein - yeas  
hypothetical prote  
serine proteinase  
hypothetical coile  
p115 protein - Myc  
hypothetical prote  
hypothetical prote  
actin-interacting

30 168.5 4.5 1115 2 T41342  
31 168 4.5 2178 2 S55805  
32 167.5 4.4 624 2 PC6003  
33 167.5 4.4 853 2 G90559  
34 167.5 4.4 1802 2 S52611  
35 166 4.4 1276 2 S11455  
36 165.5 4.4 786 2 T18469  
37 165 4.4 1465 2 S31262  
38 165 4.4 1803 2 S56894  
39 165 4.4 3724 2 T18427  
40 164 4.3 769 2 F89870  
41 164 4.3 1957 2 T38077  
42 163.5 4.3 979 2 JQ0894  
43 163 4.3 1272 2 C90593  
44 163 4.3 3216 2 C90538  
45 162 4.3 1033 2 T37715

QY 61 QSAIWSGFIKVKKSDEYTFATSGADNHVTHWVDDQEVINKASNSNKIRLEKGLYQIKIQ 120  
DB 90 QSAIWSGFIKVKKSDEYTFATSGADNHVTHWVDDQEVINKASNSNKIRLEKGLYQIKIQ 149  
QY 121 QRENPEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTSGPTVPDRDN 180  
DB 150 QRENPEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKSTSGPTVPDRDN 209  
QY 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASTASDPYDFEKT 240  
DB 210 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASTASDPYDFEKT 269  
QY 241 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTONTDSRTIISKNSTSRHT 300  
DB 270 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTONTDSRTIISKNSTSRHT 329  
QY 301 SEVHGNAEYHAFDFIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEETMGLNTADTARL 360  
DB 330 SEVHGNAEYHAFDFIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEETMGLNTADTARL 389  
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
DB 390 NANIRVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSOILAPNNYPSKNLAPIA 449  
QY 421 LNAQKDSATPTIMYNYQFLEKTKQLRLDQVYGNITATYNFENGVRVDTGNSWSEV 480  
DB 450 LNAQKDSATPTIMYNYQFLEKTKQLRLDQVYGNITATYNFENGVRVDTGNSWSEV 509  
QY 481 LPOIQETTARIIFNGKDLNVERRIAAVNPSPLETKPDMTLKEALKIAPGFPNGNL 540  
DB 510 LPOIQETTARIIFNGKDLNVERRIAAVNPSPLETKPDMTLKEALKIAPGFPNGNL 569  
QY 541 QYQKDIPTDFPNDQOTSONIKNQLAELNATNIYVLDKI KLNKMNILIRDKRFHYDR 600  
DB 570 QYQKDIPTDFPNDQOTSONIKNQLAELNATNIYVLDKI KLNKMNILIRDKRFHYDR 629  
QY 601 NNTAVGADESVEKAEHREVNSTEGLLNIDKDIRKILSGYVIEIETEGEKEVINRY 660  
DB 630 NNTAVGADESVEKAEHREVNSTEGLLNIDKDIRKILSGYVIEIETEGEKEVINRY 689  
QY 661 DMLNISLRQDGTFFDFKYNKDKPLYSINPNKYVNVAVTKENTIINPSENGDTSTNG 720  
DB 690 DMLNISLRQDGTFFDFKYNKDKPLYSINPNKYVNVAVTKENTIINPSENGDTSTNG 749  
QY 721 IKKILIPSKGYEIG 735  
DB 750 IKKILIPSKGYEIG 764

RESULT 2  
I40862  
iota toxin component Ib - Clostridium perfringens  
C/Species: Clostridium perfringens  
C/Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Oct-1999  
C/Accession: I40862; S42774  
R/Perelle, S.; Gilbert, M.; Boquet, P.; Popoff, M.R.  
Infect Immun 61, 5147-5156, 1993  
A/Title: Characterization of Clostridium perfringens iota-toxin genes and expression in  
A/Reference number: I40861; MUID: 94041637; PMID: 8225592  
A/Accession: I40862  
A/Status: preliminary; translated from GB/EMBL/DDJ  
A/Molecule type: DNA  
A/Residues: 1-875 <RES>  
A/Cross-references: EMBL:X73562; NID:g929031; PIDN:CRA51960.1; PID:g414655

Query Match 23.0%; Score 868; DB 2; Length 875;  
Best Local Similarity 31.1%; Pred. No. 2.1e-36;  
Matches 252; Conservative 130; Mismatches 279; Indels 150; Gaps 31;

QY 1 EVKQENLLNESSESSOGLGYFSDNLPQAPMVVTSSTGDSIPSSSELENIPSS-NQY 59  
DB 35 DTNQKEITNENTLSSNGLMGYFADHFKDLMLAPIKNGDLKPFEEKVKDLITEDSS 94

QY 60 FQSAIWSGFIKVKKSDEYTFATSGADNHVTHWVDDQEVINKASNSNKIRLEKGLYQIKIQ 119  
DB 95 IKIRMTGTIIIPSEDEGEYILSTDR-NDVLMQINAKGDIK--TLKVNKKGOAYNIRIE 150  
QY 120 QREN-----PTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRKST 170  
DB 151 IQDKNLGSDNLVSP-----KLYW-ELNGKNTVPIPEENLPEDYSKIDEND----- 195  
QY 171 AGTVP-----DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNTH 211  
DB 196 --PFIPIPNPFFVYRFFSAWEDEDLTDNDNIPDAYEKNGYTI---KDSIAVKWDSFA 249  
QY 212 EKGLTKYKSSPEKWSASTASDPYDFEKTGRIDKNVSPKARHPLVAAPYVHVDMENIIL 271  
DB 250 E-QYKYYVSSYLESTAGDPYTDYKASGSDKA-KLEARDEPLVAAPYVGVGMENLII 308  
QY 272 SKNEDQSTONTDSRTIISKNSTSRHTSEVHGNAEYHAFDFIGGSVAGFSNSNST 331  
DB 309 STNEHASS---DQGTVSRAATTNSKTDANTV-----GVSISAGYQNGFTGN 351  
QY 332 VADHS-----LSLAGERTWAEETMGLNTADTARLNANIRVNTGTAPIYVNLPTTSL 383  
DB 352 ITTSYHTTNDNSTAVQDSNGESWNTGLSINKGSAYINANVRVYNTGTAPYKVTPTNL 411  
QY 384 VLGNKQTLATIKADENQLSOILAPNNYPSKNLAPIALNAQKDSATPTIMYNYQFLE 443  
DB 412 VL-DGETLATIKADQNGIGNNLSFNETYPKKGLSPLALNTMDQFNARLIPINVDQLKGLD 470  
QY 444 KTKQLRLDQVYGNITATYNFENGVRVDTGNSWSEVLPQIQETTARIIFNGKDLNVER 503  
DB 471 SGQIKUETTVQSGNYGTYK-SQGI-ITEGNSWNSYISQIDSVSASIID-TGSOTFER 527  
QY 504 RIAAVNPSPLETKPDMTLKEALKIAPGFPNGNLQY-QGKDIPT-DFNFDQOTSQ 560  
DB 528 RVAAKEQGNPEDKT-PEITIGEAIKKAFSATK-NGELLYFNGIPIDESCVELLIFDDNTSE 585  
QY 561 NIKNQLAELNATNIYVLDKI KLNKMNILIRDKRF--HYDR-NNTAVGADESVEKAEHR 617  
DB 586 IIEQQLKYLDDKKLYNV-----KLERGNMILIKVPSYFTNDEFENPP--ASWNIDTKNQ 639  
QY 618 EVINSSTEGT-----LLNIDKDIRKILSGY-----VIEIETEGEKEVIN 658  
DB 640 DGLQSVANKLSGETKIIIPMSKLPKYKYVFSGVSKDPSTSNITVNIKSKEQKTDYLPV 699  
QY 659 RYDMLNIS-----SLRODGTFFDFKYNKDKPLYSINPNKYVNVAVTKENTIINP 696  
DB 700 EKDYTKFSYEFETGKDSSEIETLTSSSGVIFLENLSITELNSTPEILKEPIKVPSDQE 759  
QY 697 -----NVYAVTKENTIINPSENGDTSTNGI 721  
DB 760 ILDAHKKYADIKLDT-----NTGNTYIDGI 785

RESULT 3  
I39933  
cryptic protein - Bacillus anthracis  
C/Species: Bacillus anthracis  
C/Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 15-Oct-1999  
C/Accession: I39933  
R/Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppa, S.H.; Schmidt, J.J.  
Gene 69, 287-300, 1988  
A/Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr  
A/Reference number: I39933; MUID: 89172073; PMID: 3148491  
A/Accession: I39933  
A/Status: preliminary; translated from GB/EMBL/DDJ  
A/Molecule type: DNA  
A/Residues: 1-192 <RES>  
A/Cross-references: GB:M22589; NID:g143280; PIDN:AAA22636.1; PID:g143281

Query Match 6.2%; Score 235.5; DB 2; Length 192;  
Best Local Similarity 34.6%; Pred. No. 2.4e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;



```
QY 587 MNILIRDKRPHYDRNNAIVGADESUVVKEAHRVINSSTEGLLNIDKIRKILSGYIVEI 646
Db 1 MNILVRDP-YHYDNNNGNIVGVDDSYLKNAYKQILNWSDDGVSINLDEVDNQALSGYMLQI 59
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDFKYNDKLPYISNPN 693
Db 60 KKPNSHLNTPVTITLAGKDSGVGELYRVL-----DGAGFLDNKFDENWRSIV-DPG 112
QY 694 YKVNVAVTKEN-TIINPSENGDTSNGIKKILIFSKGYEI 734
Db 113 DDVYVAVTKEDFNATVDRENGNIA-NKLKNTLVLSGKIKEI 153

RESULT 4
G59104
hypothetical protein pX01-111 - Bacillus anthracis virulence plasmid pX01
C:Species: Bacillus anthracis
C>Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000
C:Accession: G59104
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler,
J. Bacteriol. 181, 6509-6515, 1999
A:Title: Sequence and organization of pX01, the large Bacillus anthracis plasmid harbored
A:Reference number: A59091; M01D:19445483; PMID:10515943
A:Accession: G59104
A:Status: preliminary
A:Molecule type: DNA
A:Cross-references: 1-204 <OKI>
A:Experimental source: strain Sterne
A:Note: similar to hypothetical protein in the protective antigen domain; ypa, plasmid p
C:Genetics:
A:Gene: pX01-111
A:Genome: plasmid

Query Match 6.2%; Score 235.5; DB 2; Length 204;
Best Local Similarity 34.6%; Pred. No. 2.6e-05;
Matches 56; Conservative 36; Mismatches 4; Indels 23; Gaps 6;

QY 587 MNILIRDKRPHYDRNNAIVGADESUVVKEAHRVINSSTEGLLNIDKIRKILSGYIVEI 646
Db 1 MNILVRDP-YHYDNNNGNIVGVDDSYLKNAYKQILNWSDDGVSINLDEVDNQALSGYMLQI 59
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDFKYNDKLPYISNPN 693
Db 60 KKPNSHLNTPVTITLAGKDSGVGELYRVL-----DGTGFLDNKFDENWRSIV-DPG 112
QY 694 YKVNVAVTKEN-TIINPSENGDTSNGIKKILIFSKGYEI 734
Db 113 DDVYVAVTKEDFNATVDRENGNIA-NKLKNTLVLSGKIKEI 153

RESULT 5
F82885
hypothetical protein UV482 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82885
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, B.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: F82885
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-468 <GLA>
A:Cross-references: GB:AE002145; GB:AF222894; NID:G6699476; PIDN:AAF30894.1; GSPDB:GN001
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: UV482
A:Genetic code: SGC3

Query Match 5.4%; Score 202; DB 2; Length 4688;
Best Local Similarity 21.3%; Pred. No. 0.13;
```

```
Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;
QY 19 LLGYFSDLNQFQPMVVTSTTODLSIPSELENIPSENQYFQSAIWGSFIVKPKSDYVT 78
Db 3699 LVDVYLD-NIHQNIETRKIFKDHNV-SKETINPGVTWISKHGKNWSPDTDTANFEFK 3756
QY 79 FATSADNHVTMWVDDQEVINKASNSKIRLEKRLVQIKIQYQRENPTKGLDFKLYWTD 138
Db 3757 IETQ-----DDNDVLNNIDATVFKDHNHNIQKIVRIKEN-----ND 3795
QY 139 SQNKKEVISEDNLQLPKQKSSN-----SRKKRSTAGPTVDPDR----- 179
Db 3796 WLIKQI---DNLN-PETKYKLENIELSKPLKTHHTNLSVINDKENISLITETGNPVLKV 3851
QY 180-----NGIDPSLEVEGYTDVKNKRFELSPWISNIEHKKGLTKYKSPKKNWSTASPYSD 235
Db 3852 IQQNDTIINDTQGTINVTLSGVNSK-YNGRQIKQVYKDNNNVIYESS---LITLQKGN 3907
QY 236 FEKVTGRIDQVSPPEARHPLVAAYPVHVDMENIILSKN-EDOSTQNTDSET-RTSKNT 293
Db 3908 YQLLSNLNSN-----REYREKIEIHHSNTNNFEDLEKLVNSNFIOTKNT 3957
QY 294 STGRTHTS-EVHGNAEVRHAF-----FDTGGSVSAGFS-----NSNS---STVAID 335
Db 3958 TVQWMDSSATIVTRGVNFNFKIKSEDKILENNQVVAWFAPKETIRDTNTWLQYTRPLX 4017
QY 336 HSLSLAGERWAEWTGMLNT---ADTARLANATRYVNTGTAPIYVNLPTTSLVLGKQOTLA 392
Db 4018 DVTSDPKEGTWADLNSVNSVNFKEETTKLVQFVNKPTKAKNNINNSNVILDTNISI 4077
QY 393-----TIKADENQLSOLAPNNYPSKNLAPALN-AQKASSTPITWNY--NQPLELE 443
Db 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTINFTLSGVKKSVMGKKIKLSYKSNDSI 4137
QY 444 KTQQLRLDDQVYGNATVNFENGR-----VRVDGSGNWSVLPQIQETTARIENGKD 497
Db 4138 HTNEVIESNKTQYINLLNKLKRNRYTLDVKLIDNNNVSDFPKEGNUTNSFITTRTSA 4197
QY 498 LNLVERRIAAVNSDPLETT-----KPDMLTKEA-----IKIAGFNEPNNGNQ 541
Db 4198 INVLTIEISNRASNTLKSIIKINLNDPQVLRDKDQATIVVGNKNGQAMGFITVSGNIK 4257
QY 542 YQKDKITFEDFNDQOTSQNIK-NQALNATNIYVLDKIKLNAKQNTILIRDKRHYD- 599
Db 4258 YLTATLVDLNFN-DKVINIVNISFNPKPSIAAEN-----IGDKSNII-----YNNDS 4304
QY 600-----RNIIAVGA---DSVVKEAHRVINSSTEGLLNIDKIRKILSGYIVEIEDTE 650
Db 4305 IPKLEINNDIIVNGPINKETIVVKNQKQ--NNIDVDLGLQINPKIAHNL-FLAKPKSTN 4361
QY 651 GLKEVIND--RYDMLNLSL-RODGKTFIDF-----KKY----- 681
Db 4362-----NDIETVINGSSLVWNDGKTSIRFTLNNLKANKLYSLVDVYVYLVNNSNTIVE 4415
QY 682 NDKLPLYISNPNYKAVYAVTKENTINPSENGDTS 717
Db 4416 SNKLP-KLNNINYOIK---INKSHTII--SKNGEWS 4445

RESULT 6
T28676
rhostry protein - Plasmodium yoelii (fragment)
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28676; A45521
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.
Mol. Biochem. Parasitol. 76, 329-332, 1996
A:Title: Comparison of two members of a multigene family coding for high-molecular mass
A:Reference number: Z20507; MUID:97077455; PMID:8920022
A:Accession: T28676
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2401 <SIN>
```

A;Cross-references: EMBL:U36927; NID:gl041784; PID:gl041785; PIDN:AB41263.1  
R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-246, 1990  
A;Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co  
A;Reference number: A45521; MUID:91101660; PMID:2270106  
A;Accession: A45521  
A;Status: Preliminary  
A;Molecule type: DNA  
A;Residues: 2260-2401 <KEP>  
A;Cross-references: GB:M34281

Query Match 5.1%; Score 190.5; DB 2; Length 2401;  
Best Local Similarity 21.7%; Pred. No. 0.19;  
Matches 182; Conservative 119; Mismatches 316; Indels 221; Gaps 43;  
QY 1 EYKQENRLINSESSQGLGYFSDLN-----PQAPVVTSSITGDLST 45  
DB 402 EYAKENQVNVYKSNLEIKHYNDOINIDNIKEAKQYNDQFKHEKMTIPPNEKMYQK 461  
QY 46 PSELENIPSENOYQSAI--WSGPIKV---KKSDEYTFATSAOHVMTWDDQEV---I 97  
DB 462 PSIEIKMKDE--FLSKVKNYDFDKVYKEVSEHNKFTLTNKKITEVSDSEIKYE 518  
QY 98 NKAASNKIRLEKGLYQIKIYQRENTEKGLDFKLYWTSQNKKEVISSDNQLPELK 157  
DB 519 NKFNDSKSLNETKK--SIEEYQNLTKKYDD---YIKVCLNTNELITNCHNKQTLK 573  
QY 158 QKSSNRK--KRSTGAGTPVPRDNDGIPD-----SLEVEGYTVDVKNKRTFLSP 205  
DB 574 DKLQNIKTIKETNIDKTYTDKFNILTDKTELETKTGLSLNHNESNKELLTFYD 633  
QY 206 WTSNHEKGLTKYKSSPKWSTADSPYDFKVTGRIDKNVSPPEARHPLVAAPVHVD 265  
DB 634 LKANLGNKENMLYQFNEK---SKAVEDIKKQVNDINKIVSN-----IBIT 677  
QY 266 MENILSKNEDQSTQNTDSETRTISKNTSTSTHTSEVHGAEVHASFDDIGSVSAGFS 325  
DB 678 IYTSIYNINED---TENE---IGKSELNTKVLK-----KVKANVT 713  
QY 326 NNSSTVAI--DHSLSLAGERTWAETWGLNTADTARLANIRYVGTAPIYVLTSLV 384  
DB 714 NLNEIKELKDYDFQFGK-----EKNIKYPDEN--KIKNDIDTLNOK 754  
QY 385 LGKN-QTLATIKAD-ENQLSQI-----LAPN---NYYP---SKNLAPALNAOK-- 425  
DB 755 IDKSIETLTETIKNSNHEIDKQIDKUKVFNKTMFNDPKIEKKIENIVEKIDKKK 814  
QY 426 -----DASSTPIYMNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGSMNSE 479  
DB 815 NIYKEIDKLLNE-SKENDKTSLEKLNINLSYKSLGNLPLQOIDEEKKAEHTIKAME 874  
QY 480 V-----LPOIQTETARIIFNGKDLN-----VERIAAVNPSDP-----LET 518  
DB 875 AYIDDLNKKKSQEI---EKEMNINMDIKMDIHKEMKALNISHDDYKIYHTTSKHBEK 931  
QY 519 PMTLKEALKIAPGFNEP---NGNLQYQKDKITEFDNFDQOTSQNIKNQLAELNATNIY 575  
DB 932 ISPIRKNLSKLIQDFSEESYINDIKKELEKNVLE-----SQNNNTDINQYLSKIE--NIY 984  
QY 576 TVLDKTKLNAKQNLIRDKRF--HYDRNIAVGADESVEVKEAHREVINSGTEGLLNIDK 633  
DB 985 NIL---KLANKIKIIDKVKREYTDIEKKN-----KKINAELSNS-----1020  
QY 634 DIRKILSGYIVETEDTEGLKE-----VINDRY-----DMLNSSLRQDQKTFIDPKKYND 683  
DB 1021 --EKI---ITQKENSLSKECOSKIKSTIDDNVYSECIKNTNL-----KTVIVNEKN--1069  
QY 684 KLPYISN-PNYKVV-----YAVTKENTIINPSNGDSTN--GKKILILFSKK 730  
DB 1070 -INTYFNAEYQNQVNSLNFNNT-EMADTKSQYLINIKKNGTNTDYNIKELKEHKKK 1126

RESULT 7

B64635  
toxin-like outer membrane protein HF0922 - Helicobacter pylori (strain 26695)  
C;Species: Helicobacter pylori  
C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C;Accession: B64635  
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, I  
Nature 388, 539-547, 1997  
A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C  
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A;Reference number: A64520; MUID:97394467; PMID:9252185  
A;Accession: B64635  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-2529 <TOW>  
A;Cross-references: GB:AE000602; GB:AE000511; NID:g2314060; PIDN:AD07969.1; PID:g23140

Query Match 5.0%; Score 189.5; DB 2; Length 2529;  
Best Local Similarity 21.0%; Pred. No. 0.23;  
Matches 158; Conservative 92; Mismatches 264; Indels 239; Gaps 35;  
QY 123 ENTEKGLDFKLYWTSQNK---KEVISSDNQLPELKQKSSNRKSTKSAGTPVPDR 178  
DB 729 QNPASVSWGVTLW--QNKTSNKGIIYID-----PNLSGQSGQSGNTLSTYANLF--- 779  
QY 179 DNDGIPDSLEVEGYTVDVKNKRTFLSP---WI-----SNHEKKGLTKYK 220  
DB 780 ---GRSFSVNIQNTGLIIGNNTESVNSGLIWIHGFGYITGTFSAAIY---LTNMF 832  
QY 221 SSEKYST-----ASD-----PYSDFEKVTGRIDKNVSPPEARHPLVAAPIV 262  
DB 833 KTGEVNSDDGGANTTFKASDNITWGLNYNDAETVTKIOTGAS---CHSAYTFDALN 889  
QY 263 HVDENIILSKNEDQSTQNTDSETRTISKNTSTSTHTSEVHGAEVHASFDDIGSVSA 322  
DB 890 NISVTNSSF-----DMTWGKFSFSAKNISFS-----NASF-----S 921  
QY 323 GFSNSNSTVA---IDHSLSLAGERTWAETWGLNTADTARLANIRYVGTAPIYVNL- 378  
DB 922 GFTNPGQSSVISANATNSLSFNSR-----LNGAVYNLQANSLIFF-NTQAVFNVL 973  
QY 379 -----PTSLVLGKNQTLATIKADENQLSQIILAPNNYPSKNLAPIAL-NAQKASS 429  
DB 974 SRGTSNPNATQLLGNNTFTLSS-----QSLNFGDITLONNANITLGNKSQA 1025  
QY 430 TPTMNVNPOLEKTKQLRLDQVYGNIAFYNFENGVRVDTGSGNWSEVLPOIQTET 489  
DB 1026 NSULTDNNSLSLQNSVLANNTSAFNQASLIYNGS-----QATFN 1069  
QY 490 RIIFNGKDLNL-VERIAAVNPSDPLETTKPDMTLKEALKIAPGFNEPNGNLQYQKDKIT 548  
DB 1070 SLFPNGTSLNASSKLNASNASFSNNTT---INLDDSVLSASNTSSLANINFQASQA 1126  
QY 549 EF-----DFNFDQOTSQNIKNOLA-----EL 569  
DB 1127 DFGGNTIIDTASTNFDSSSLNFNLTANGALNFNGYTPSLTKALMSVSGQFVLGNNGDI 1186  
QY 570 NATNIYTVLDKILNAKQNLIRDKRF-----HYDRNIAV 606  
DB 1187 NLSDI-NIFNITKSVTYNIIILNAQKGITGSGANGYKEKILFYGMKTKQNTATSDNNIQTW 1245  
QY 607 A-----DESVVKEAHR-----EVIN--SSTEGLLNIDKDI-----RKILSGYIVEI 646  
DB 1246 SFINPLNSSQIIQESIKNGDLTIEVLNPNASNTIFNIAPELYNQASKQNPYTSYDY 1305  
QY 647 EDTGLKEVINDRYDMLNLSL-----RODKTFIDPKKYNDKLPY---ISNPNYK 695  
DB 1306 SDNOA-----GTYLTSNIKGLFTPKGQTPQAPGYPFPNPQLSSLNINPKGFSSENK 1360  
QY 696 VNVAVTKENTIIN--PSENGDSTNGIKKIL 725  
DB 1361 TLLGILSQNSATLKEMIESQNLDNITN-INEVL 1392



382 SLVGLKQNTLTIKADENQSLQILAPNNYPSKNLAPIA---LNAQKASSTPITMNYN 437  
1257 GQVVTGEAVTPSVI---DNILSKI---ENEVEVLKPLAGYRSLKKQLNENVMVTNVN 1310  
438 ---OFLKLEKTKQLRLDQV-YGNIAT-----YNFENGRVRVDTGNSWSEVL 481  
1311 VKDIILNSFNKENFNKV-LESGLIPYKDLTSSNNVVKDPYKFLNKEKRDKFLSSYNVIK 1369  
482 POIQETTARIIPNGKDLNLVERRIAAVNSDPL-----ETTKPDM-TLKEALKIAGF 533  
1370 DSID-----TDINFA-----NDVLGYKYLSEKYSDDLISIKKYINDKQGE 1410  
534 NEPN---GNLOVQKQDIPE-----FDNFDDQOTSQNIKNQLAELNATNIYT 576  
1411 NEKILPFLNNIETLYKTVNDKIDLFVHLEAKVLNHYEK---SNVEVKIKELN--YLKT 1465  
577 VLDKIKLNAAMN--ILIRDKRPHYDRNN-----IAGV-ADSVVKEAHEVINSSTEGLL 628  
1466 IQDKLADFKNNSNFVGIADLTDYNNHLLTKFLSTGMVFENLAKTVLSNLLDGNLQW-M 1524  
629 IATDKD--IRKILSGYIVIEDTEGLKEVINDRYDMLNITSSLRQDGKTFIDPKYNDKLP 686  
1525 LNIHQHCVK---OCPQNSGCFRHLDE-----RECKCLNLYKQSGDKC- 1566  
687 LYISNPNYKVNVAVTKENTIIINPSENG-----DSTNGIKKI 724  
1567 --VENENPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1599

RESULT 10  
B89921  
hypoetical protein ebba [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: B89921  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud  
ma, A.; Mizutani-Jui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; MUID:21311952; PMID:11418146  
A:Accession: B89921  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-6713 <KUR>  
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BAB42527.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: ebba

Query Match 4.9%; Score 183; DB 2; Length 6713;  
Best Local Similarity 20.4%; Pred. No. 2.1; Indels 256; Gaps 43;  
Matches 172; Conservative 125; Mismatches 125

9 LNESSSSQGLLYFSDNFQAPMVVTS-----STTGDLISIPSELENIPSENQYFQSA 63  
2468 VRQAKSDAKANLG-TLTHLNAQKQDLTSQIEGATTGVNSVTKAQDLQDGMQRLESA 2526

64 IWSGFIKVKSDVEY-----TATADNHVTWVDDQVINKASNSKIRLEKGLYQIKI 118  
2527 I-ANKDQTKASENYIDAPTKTAPDNAT---QAESYLNKHGNTKDK-----QAVEQ 2576

119 QVRENPTKGL--DFKLYWTDSONKKEVISSDLQLPE---LKQKSNRKKKSTSGAP 173  
2577 AIQSVTSTENALNGDANIQCATEATQAIIDLNTQLNTPQTKALQOVNAAQR----- 2628

174 TVPDRNDGIPDSLEVEGYTVVQKRTFLSPWISNIHEK-----KGLTKYKSSPEK 225  
2629 -----VSGVT-DLKNASATSLNAMDQLKQATGDHDTIVAGNNTYVAGNPNASPDK 2672

226 WSTASDPYSDPEKVTGRIDKNVSPBARPLVAAPVIVHVMENITLSKNEDQSTQNTDSE 285  
2673 QGAYTDYNAAKNIYNG--SPNVTNADVTATQV-----NNAETSLINGDTN 2719

286 TRTI---SKNTSTSRKTHSEVHGNAEVHASFPDIDGYSVA-----GFNSNSNSTVAIDHS 337  
2720 LATAKQAKDALRQMTLSDAQKQS-----ITGIDSAQTQVTGVQSVKDNATNLDNA 2771  
338 L-----SLAGERTWASTMGLNTADTARLANRANRYVNTGTAPIYVNLPTS----- 382  
2772 MNQLRNSIANKDEVKASQPYVDADTKQNA-----YNTAVTSAENIINATISQPTLPSAVT 2827  
383 -----LVLGKQNTLATIK---ADENOLSQILAPNNYPSKNLAPIALNAQKQAS 428  
2828 QAAQNVNTKALTALGAQNLANKQETTANIRLSHL-----DQV-----YQNI 459  
429 STPITMNYN-QFLELEKTKQLRL-----DQV-----YQNI 459  
2872 NTQVTNAPNISTVNVQVTKRAEQLDQAMERLINGIQDKQKQSVNFTDADPEKQTYANNA 2931  
460 AT-----YNFENGRVRVDTGNSWSEVLPOIQE--TTARIIFNGKDLNLVERRI--AAVNPS 511  
2932 VTAENIINOANG-----TNANQSVYEAALSTVTTTKQALNG-----DRKVTDAKNAN 2980  
512 DPLET-----TKPDMTLKEALKIAPGFNPNQNLQ--YQKQDIIE 549  
2981 QTLSTLDNLNNAQKAVTGNGINOAHTVAEVT--CAIQTAQELNTAMGNLKNLNDKDTTL 3038  
550 FDNFED-----QQTSONIKNQLAELNATNIYTVLDKIKLNAKQMLILIRDKRPHYD 599  
3039 GSQNFADADPEKKVAYNEAVRVAENILKSTGTNV---PKDQVEAAMN-----QVN 3086  
600 RNNTIAGVADSVVKEAHEVINSSTEG--LLNIDDKIRKILSGYIVIEDTEGLKEVIN 657  
3087 TTKAALNGTQNL--EKAKQHANTADIGLSHLTNAQKALQVQSTTVAEAQNSQKAN 3144  
658 DRYDMLNIS---LRQ---DGKTFIDFKYNDKLPYISNPNYKVNIV--AVTKENTII 708  
3145 -----NVDAAMDKLRQSIADNATKQNYTD-----ASPN-KKDAYNAVTTAQGII 3191  
709 NPSEN 713  
3192 DQTTN 3196

RESULT 11  
JC6009  
surface-located membrane protein lmp3 precursor - Mycoplasma hominis  
C:Species: Mycoplasma hominis  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: JC6009  
R:Radefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.  
J. Bacteriol. 178, 2775-2784, 1996  
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene syate  
A:Reference number: JC6009; MUID:96213016; PMID:8631664  
A:Accession: JC6009  
A:Molecule type: DNA  
A:Residues: 1-1302 <LAD>  
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336  
C:Genetics:  
A:Gene: lmp3  
A:Genetic code: SGCG  
C:Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homology  
C:Keywords: duplication; membrane protein  
F;24/Domain: signal sequence #status predicted <SIG>  
F;25-1302/Product: surface-located membrane protein lmp3 #status predicted <MAT>  
F;957-992/Domain: tetratricopeptide repeat homology <T1>  
F;993-1026/Domain: tetratricopeptide repeat homology <T2>  
F;1089-1120/Domain: tetratricopeptide repeat homology <T3>  
F;1154-1190/Domain: tetratricopeptide repeat homology <T4>

Query Match 4.8%; Score 181; DB 1; Length 1302;  
Best Local Similarity 17.9%; Pred. No. 0.23;  
Matches 137; Conservative 141; Mismatches 285; Indels 204; Gaps 28;

1 EVKQENRLNSESQGLLYFSDNFQAPMVVTSSTTGDLISIPSELENIPSENQYF 60

Db 476 QVDEANKSIKQALNALIDKANTLLPQNDNDSEIVKAKESLNABEITWAKAVNQNDNASM 535  
Qy 61 QSAIWSGFIKVKSDSEYTFATSAADNHVTWVDDQEVINKASNNKIRLEKGR--LYQIKI 118  
Db 536 QSA-----KSSLDKVKTIQNLTEFNKDKDAKFELEQTRKIDNFLT 579  
Qy 119 QYQRENPTKGI-----DFKLYWTDSQNKKEVISSDNLQLPELKQKSSNSR--KKRS 168  
Db 580 DDVKNPNYATLVKDLTWAKDDKSVTKSSNKSIIAAND-----ELQALDKAKVAKQOI 635  
Qy 169 TSAGTVPDRNDGIPDSLEVEGYTV---DVKNKRTFLSPWISNIHEKGL----- 216  
Db 636 DEANKSIKQELSDSITNANQLLNKLVDSKDIOAKATELSOEIQSASQELMNNPTSMQS 695  
Qy 217 -----TKYKSSPEKWSIASD--PYSDFEKVTGRIDKNVSPPEARPLVAAPIVHVDM 266  
Db 696 AKESLDKAVTEITKLETFNKKDVKFKELEKTRKIDDEFINTKNTNP----- 743  
Qy 267 ENILSKNEDOSTQNTDSETRTISKNTSTSRHTSEVHGNAEVAHSPFDIGGSVSAQFSN 326  
Db 744 -----NYSTLSELTSEKSDSKSVTSSNKSIDI----- 771  
Qy 327 SNSSTVAIDHSLUSLAG-ERTWAETMGLNTADTARLANIRVYNTGTAPIYVLTSTSLVL 385  
Db 772 -ETANTEKQALAKANTDKAQNDAKSTKE--QLNNSISSANTLLAKLTD----- 819  
Qy 386 GKNOTLATIKADENOLSQLAPNNVYPGKNLAPI-ALNAQKQDASSTPI-----TWNYNQF 439  
Db 820 -KDNVIOQAKTELEK--EVQKANOAVASNTASMQSAKSLDAKYTEITKLETFNKKDKD 876  
Qy 440 LELEKTKQRLTDQVYGNIAATYFENGVRVDTGSNWSEVLPOIQTETARIIIFNGKDLN 499  
Db 877 VKFKELEQTRKIDDEF-----INTKNTNPDYXSTLSELTSEK----- 912  
Qy 500 LVERRTAAVNPDPLETTKPDMTLKEALKIAPGFNEPNGNLQYQCKDITDFENFDQOTS 559  
Db 913 -RDKNSIITNSNKSIDETANTEKQALAKA-----NTDKQAD-NLARSTK 957  
Qy 560 QNKKQALNATNIYTVLDKTKLNKAKMILIRKFRHYDRNNIAVGAD-----ESVVEKAH 616  
Db 958 EQLNKSISSAN-----TLAKLT-----DKNTTIOQAKTELEKEVQKAN 996  
Qy 617 REVINSSTGLLNIDKDIRKILSGVIVEIDTEGKEVINDR-YDMLNIISSLRDQKTF 675  
Db 997 QAVASNTASM-----QSAKSSLDKAVTEI--TKLETFTNKDKDVKFELEQTRKIDDEF 1049  
Qy 676 IDPKKYNDKPLVYISNPKNVYAVT-----KENTIIINPSENGDTST 718  
Db 1050 INTNK-----TWPNYSTLISELTSEKSDSKSVTSSNKSIDET 1087  
RESULT 12  
T28677  
rhoptry protein - Plasmodium yoelii  
C:Species: Plasmodium yoelii  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T28677; C45521  
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 65, 171-177, 1994  
A:Title: A gene coding for a high molecular mass rhoptry protein of Plasmodium yoelii.  
A:Reference number: Z20508; MUID:95021522; PMID:7935623  
A:Accession: T28677  
A>Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-2269 <KEE>  
A:Cross-references: EMBL:L27838; NID:9457145; PID:9457146; PIDN:AAA21304.1  
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-246, 1990  
A:Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co  
A:Reference number: A45521; MUID:91101660; PMID:2270106  
A:Accession: C45521  
A>Status: preliminary

A:Molecule type: DNA  
A:Residues: 2131-2269 <KE2>  
A:Cross-references: GB:M34283  
Query Match 4.8%; Score 181; DB 2; Length 2269;  
Best Local Similarity 20.5%; Pred. No. 0.53;  
Matches 161; Conservative 121; Mismatches 301; Indels 204; Gaps 35;  
Qy 1 EVKQENRLNSESSESSQGLLYVYFSDNFPQAPWVVT---SSTTGLSLIPSS------L 50  
Db 587 KVENVTNINLK-K--EKLKHVDFDFGKEGNIKYTDKIKKINDDIMAVSQIIDQHINGL 643  
Qy 51 ENIPSNQVQSAIWSGFIKVKSDSEYTFATSAADNHVTWVDDQEVINKASNNKIRLEK 110  
Db 644 DDIQKSESIVSEMKQEQINKLEKVSNTETI--SNDNVEGIGKKQOIIVTKIDKKKNIIYEI 701  
Qy 111 GRLYQIKIQYQRENPT-EGLDLFLYWTDSQNKKEVISSDNLQLPELKQKSSNSRKRST 169  
Db 702 NKLSISIKETKNTSLEKVKDINLSY--GQNLGNLFLE--QIDBEKKAENTIK---- 752  
Qy 170 SAGTVPDRDN-----DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPE 224  
Db 753 SMEAYIDDLNKKKQSEIETEMDIK---MDINKEMEALK--ISHDDDKKCHDKSKNHKE 807  
Qy 225 KMTASDP-----YSDFEKVTGRIDKNVSPPEARPLVAAPIVHVDMENIILSKN 274  
Db 808 NISDIYDKSKIIQDPSRESNDIINKLQKNVSESNH-----NSDINQCL---N 855  
Qy 275 EDOSTQNTDSETRTISKNTSTSRHTSEVHGNAEVAHSPFDIGGSVSAQFSNSSTVAI 334  
Db 856 EVANIYNI-LKLNKIKIIDKVKYETSEIEKNK-----NINDELNNSEKVIKKI 904  
Qy 335 DHSLSLAGERTWAETMGLNTADTARLANIRVYNTGTAPIYVLTSTSLVLGNQTLATI 394  
Db 905 EGLSLKECKSKIN---STLDDKDIDECIKIN-----VLKKN----- 939  
Qy 395 KADENQLSQLAFNNYPSKNLAPIALNAQKQDASSTFITMNYNQFLELEKTKQRLDQ 454  
Db 940 -----ILNEB-----TNIT---NHFKNAEB----- 956  
Qy 455 YGNIATYFENGVRVDTGSNWSEVLPOIQTETARIIIFNGKDLNVERRIAANVPSDPL 514  
Db 957 -YKNIVLSNFN-----IEMADNKSQVILEIKONGT---NDHDYNIKELK-SHKDSNGY 1007  
Qy 515 ETTKPDMTLKEALKIAPGFNEPNGNLQYQCKD-----ITEFDNFDQOTSONIKN 564  
Db 1008 K-TEADQNKAKIQK-----NKELEQYKEEVTVLLNKYAVELKKNKFD-KTKDQSKQ 1057  
Qy 565 QLAELNATNIYTVLDKTKLNKAKM-----ILIRDKFRHYDRNNIAVGADSESVVEKAH 618  
Db 1058 IIEIKDAHNYCTLESQSEKQKNEIKNEIKIHIEDEVANNNDKGNKAITSIKVSVEPPKTK 1117  
Qy 619 VIN---SSTEGLLNIDKDIRKILSGVIVEIDT-----EGLKEVINDRYDMLNI 665  
Db 1118 IIKINEIRTKSDCKLETNDLEKQISNLSIDTQETKLTENGKQLKLEEL-----L 1168  
Qy 666 SSLRQDQKTFIDFKKYNDKPLVYISN-----PNYKNVYAVTKENTIIINPSENGDT 716  
Db 1169 ESLKKQKKNIEDQKDELDEVNSKIKNIENTVNHCKKNYEIGI--VEKINEIAKTNKQIE 1226  
Qy 717 STNGIKK 723  
Db 1227 STKELIK 1233  
RESULT 13  
A86827  
Hypothetical protein yvfg [imported] - Lactococcus lactis subsp. lactis (strain ILL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: A86827  
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrh  
Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A;Reference number: A86825; MUID:21235186; PMID:11337471  
A;Accession: A86827  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1072 <STO>  
A;Cross-references: GB:AB005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146  
A;Experimental source: strain ILL403  
C;Genetics:  
A;Gene: yqfG

Query Match 4.8%; Score 179.5; DB 2; Length 1072;  
Best Local Similarity 19.7%; Pred. No. 0.21;  
Matches 164; Conservative 126; Mismatches 345; Indels 197; Gaps 32;  
QY 2 VKQENLLNESSSQGLLYFSLNFCQAPWVTSSTTGLSLPSSSELEN----- 52  
Db 300 VEQDVASSESTQDANSASLPISEASSVTDNLNLSISLSDSSISSTQTSNQSQASSTA 359  
QY 53 ----IPSENCYFOSA---TWSGFIKKKDEYTFATSDAHVHTMWYDDQEVINKASNSNKI 106  
Db 360 EISYDSENLSLSSNQINNSNSEKDSNQSSLGSSMSNSESEHSNSNINETNNSSEI 419  
QY 107 R-----LEGRLYQIKIQRENPTKGLDFKL-----YWDSQNKKEVISSDNQLPLK 157  
Db 420 TNLPPSPTEGNSVSDQTSSEASTNSNSISLSPSNISSTSDSSATNSDFFGNVAEVA 479  
QY 158 QKSSNSRKKRSTAGPTVPDRNDGI---PDSLEVEGVTVDVKNKRTFLGFWLSNTHK 213  
Db 480 NNSLASVNNSSSVLSSTSTADNLGNGSGDNLYKD--SSEISTSGAFLS--SNQTS 534  
QY 214 KGLTKYKS---SPEKWS-----TASDPVDFEKTGTRIDKNVSPEARHPLVAAPVI 261  
Db 535 EASTNSNSISLSPNISSTVLESTTSNFSNVAEVANNASLASVNNSSSVLSSTSTA 594  
QY 262 VHVDM-----ENI-----ILSKNEQDQNTDSETRTISKNTSRTHTSP-- 302  
Db 595 DNLEINQPSDLTKDSSBISITSGAFLSSNQTSSEASSNSMSSINSPSLSLSLNSESA 654  
QY 303 -----VHGNAEVAHFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMG 351  
Db 655 TQNSNSEATKVDNNSSTHSSNLSGNSDSDSDSDSDS--DSS----- 697  
QY 352 LNTADTARLANIRVYNTCTAPIYVLPPTSILVKGKNOTLAIKADENQLSQILAPNNY 411  
Db 698 -NLSGSPNLETN-QTISKPSVNNISENPKKVSSN-----SVQENSTHEMSTN--- 746  
QY 412 PSKNL-APIALNA---QKDASS-----TPITMNYNQFLEKTKQLRLDQVYG 457  
Db 747 PKSSISSP-STTSSSQKSSQSNLLNTTEGINNPITFN-NSSSENAAAIL-----TSYS 800  
QY 458 NIATYFENGREYV-----DTGSNWSEVLPOIQTETARIIFNGKDLNVERRIAAVNP 511  
Db 801 NNSSESSETGCLYISNEAQRDNGSEISHSLPS-----SNSNENNVSSIQSOAIL 851  
QY 512 DPLETTKPMWLKEALKAFGEFNPNGNLQYQKDIPEDFNFDQOTSQNIKNQALNELNA 571  
Db 852 SKGSTNKRSSSLSIINSTSHPNQEDN---QSNSSD-----EVKSNNVESILGQLNS 900  
QY 572 TNYTVLDKI---KLNAKNILIRDKRFHYDRNNIAGADESVVKEAREVINSSTEGLL 628  
Db 901 ISNKTMSNLSLTSQKLSVIYLPKSKVTKNEKNSNTVSEKLIKTPOKN-DESNLQGI 959  
QY 629 LNDKDIRKILSGVIVEIEDTEGLKEVIN-----DRYDMLNLSLRQD-GKTFIDFK 679  
Db 960 TALDLSFNK-----EVTMEDSKTVDPKVLNDENGDSQNKKTSTIAKDKNKFVKES 1012  
QY 680 KYNDKULPLYISNPYKNVYAVNTKENTINPSNGDTSNKGKILIFSKG 731  
Db 1013 EFNSKIL-----DSDNNILKTKVLLKKG 1036

RESULT 14

S68218

botulinum neurotoxin type A, nontoxic/nonhemagglutinating - Clostridium botulinum (stra  
C;Species: Clostridium botulinum  
A;Variety: strain NIH  
C;Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 02-Jun-2003  
C;Accession: S68218; S74301  
R;Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.  
FEBS Lett. 376, 41-44, 1995  
A;Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin component  
A;Reference number: S67988; MUID:96096783; PMID:8521962  
A;Accession: S68218  
A;Molecule type: DNA  
A;Residues: 1-1193 <FUJ>  
A;Cross-references: EMBL:D67030; NID:g2160224; PIDN:BAAL1050.1; PID:d1011710; PID:g1132  
A;Experimental source: strain NIH  
A;Accession: S74301  
A;Molecule type: protein  
A;Residues: 1-13/145-155 <FUI>  
A;Experimental source: strain NIH  
C;Genetics:  
A;Gene: ant  
C;Superfamily: tetanus toxin  
C;Keywords: neurotoxin  
F;1-1193/Product: botulinum neurotoxin type A, nontoxic/nonhemagglutinating #status exp

Query Match 4.7%; Score 179; DB 2; Length 1193;  
Best Local Similarity 19.6%; Pred. No. 0.26;  
Matches 181; Conservative 131; Mismatches 319; Indels 294; Gaps 46;  
QY 42 DLSIP-----SSELENIPSENCYFQSAWSGFIVK--KSDEYTFATS-----ADNHT 88  
Db 241 DLVIFRLSELENIYSQINIVDLLVSGIDPKFINTDPYFTDNYFSNAKVFEDHRN 300  
QY 89 MWYDDQEVINKASNSKIRLEK-----GRLYQIKIQY-----ORENPTKGLDF 132  
Db 301 IYETEIEGNAIGNDIKLRKQKFRININDIWELENLNYFSKFSIMMPDRFNALKHFYR 360  
QY 133 KLYW-----TDSNKKKEVSS----- 148  
Db 361 KQYIKIDYPENYSINGFVNGQINAOQLSDRQDIINKPEEINLLNGNNSVLMRSNIY 420  
QY 149 -----DNLOLP-----ELKQKSNRKRSTAGTPVDRDNDGIPDSLEVEGY 192  
Db 421 DGLKSTVDFFSYKIPYNRAVEYHFNNSDLSLNVNIGVI-----DNIPEIIDVNPY 474  
QY 193 TVD-----VANKRTFLS--PWISNTHKGLTKYKSSPEKMWSTADDPYSDEKVTG 241  
Db 475 KENCDFSPVQKITSTREINTNIPWPNY-----LQAMTNNEKFSLS-----SDFVEVVS 525  
QY 242 RIDKNVSPEARHPLVAAPYIVHVDMENII--LSKNEDQSTQNTDSE---TRTISKNTST 295  
Db 526 SKDKS-----LVYSFLSNVMFYLDSIKDNSPIDTDKKYVLWLREIFRNYSF 571  
QY 296 SRHTTSEVHGNAEVH-----ASFDDIGGSVAGFSNSNS-STVAIDHSLSLAGERT 345  
Db 572 DITATOEINTNCGINKVYTWFGKALNLTNTSDSFVEEQNLGALSINKENLSMFI 631  
QY 346 WA---ETMGLNTADTARLANIRVYNTG--TAPIYVNLPT-----TSLVLGKNQT 390  
Db 632 YELPNDMLGLPLNDLEKLFNIYSKNTAFYKLIYVFLQWMTQYYSQYFDLICMAKRSV 691  
QY 391 LA---TIKADENQLSQILAPNNYPSKNLAPIAL---NAQKDASS--TPITMN----- 435  
Db 692 LAQETLIKRIIQKLSYLIGNISSDNLALANLNTTNTLRDISNESQIANNVDSFLNN 751  
QY 436 -----YNQFLELEKTKQLRLDQVYG--NIATYFENGREYVVDTSNWSVLPQ 483  
Db 752 AATCVFESNYPKFISF-----WEQCINNINIKTEF-----IQKCTINNEDEKLQ 798  
QY 484 IQETTABRIIFNGKDLNVERRIAAVNPSPDLETTKPKDMLKEAL-----KIAFGNPNNGN 539  
Db 799 INQN-----VFNSLDPEFLN-----IQNMKLSFSSETALLIKEETWPVELVYAFKEPGNN 849

```

QY 540 -----LOYGKDI-TEDFENFD-----QOTSONIKNOELAEINATNIYTV----- 577
Db 850 VIGDASGKNTSIEY-SKDIGLYGINSDALYNGSNQISFNSNFFENGLTNSFSIYFWL 908
QY 578 ---LDKIKLNAXM-----NILRD-----KRPFY----- 598
Db 909 RNLGKDTIKSLIGSKEDNCGWIEYFQDTGLVFNMDNSNGNEKNIYLSDVNSNSWHYITI 968
QY 599 ---DR--NNIAGADESVV-KEAHREVINSSTEGLLNIDKDIRKILSYIYIEIDTGLK 653
Db 969 SVDRLUKEQLLFIIDNLVANSIKELIYNSNIISLSENNPSYIEGLTILNKPTTS-Q 1027
QY 654 EVINDRYDMLNSSLRQDKTFIDFKYNDKLPYISNPYKNV---YVAVTKENTI--- 707
Db 1028 EYLSNYFEVLANSYIRDSNEERLE---YNKTYQLY---NVVFSKPICEVKONNNIYLT 1080
QY 708 INPSENGDSTNGIKKILIFSKGY 732
Db 1081 INNTNMLNQASKFKLLSINPNKQY 1105

RESULT 15
C97033
uncharacterized protein, probably surface-located [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: C97033
R:Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C10
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97033
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1227 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK79054.1; PID:G15023995; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC1080

Query Match 4.7%; Score 179; DB 2; Length 1227;
Best Local Similarity 18.9%; Pred. No. 0.27;
Matches 166; Conservative 136; Mismatches 288; Indels 288; Gaps 44;

QY 37 SBTGDLSPSSSELENIPSENQYFQSAIW-SGFIKVKSDSEYTFATSADNHNVTWVDQOE 95
Db 14 SVTTSVALLSKPAIKAYAADNSVYKSSISNSNEINVKGE-----VQSEKE 58

QY 96 V-INKASNSKIRLEKGRLYQIKIQVORENPTEK-----GLDFKLYWTDSQNKKEVI 146
Db 59 VNVKNKNDNS-KVSSSNENQVS---NKENSFPKVSSEISQISINKNVNLQVQNNKSVL 114

QY 147 SSDNLQLPKOKSNSRKK-----RSTSGTVPDRNDGIPDSLEVEGYTVDVKNKRT 201
Db 115 AASNVDDEVKINGSNVQTSYIAIGETKVKPDTLLIINKAIVDARSAG--TDLSEVEI 172
QY 202 F-----LSPWISNI-----HEKKG-LTKYK 220
Db 173 YDIVSQTAEIQAEPFRINDGVANVSDYTLIGATFVNDANLDSVNKYFHKRYATVTKF 232

QY 221 SGPWKSTA-----SDPYSDFEKTGTGRIDKNVSPPEARHPLVAA 258
Db 233 DVATKTSNALKNNNGGGGTDYTALEVSGVTPYLDL-----VNKNIVKEKQ----- 280

QY 259 YPIVHVDMENILSKNEDOS-TQNTDSETRTISK-NTSRTHTSEVHGNAEVHASFEDI 316
Db 281 -----NKGRLTIITEISDSAASTIARINTALDN-----MDA 311

QY 317 GGSVSAGFSNSNSTVAIDHSL---SLAGERTWAETM-----GLNTADTARLNANI--- 364
Db 312 GVATLEDYQAI GANNVPQLHVADVNSLAMDQRWGVSEALDGINTIMTYINNINSVGTE 371

```

```

QY 365 -RYVNTGTAPI-----YVNLPTTSL--VLGKNQTLATIKADENQLSQILAPNNY----- 411
Db 372 DDIYINSHAVDSNEGNIIDYDILNANIIEKKTAKGQDL-TIPEVANVVKEVKTLDDFFNHAA 430
QY 412 -----PSKNIAPIALNAQKQDASSTPTMTNTNQFLELEKTKQLRLDQV----- 455
Db 431 AGOTTLQDYKNVDPNNAQVQDDVAT-----LSDMLKTRDCKTLKALQDKIDSILNSLQK 484
QY 456 ---YGNIATYFENGVRVDTGSNNWSEVLPOIQETTARI-IFNGKDLNLVERRIAAVNP 510
Db 485 INSGIGNIDDYS-----KIQTEAVDASKLEAVNDDIKIKADKGEDLTQIEIRDSVKKT 538
QY 511 SDPLETT---KPDWTLKBAKIAF-GFNEPKNLQYQKQDTEFDNFDDQOTSQNIK-- 563
Db 539 IDYNSTSVSKGDSVSDYITIGIDGVTEI--NIEFVNERIKESGITI---TIENIKV 593
QY 564 ----NOLAELN---ATNIYTVLD-----KIKLNKMMILIRDKRFHYD 599
Db 594 IEFIVQLSEVYVITGVGTIVVDYKTLGINNVNDNNIYINAEKKNKDKVKIQDITRVD 653
QY 600 R--NNI-----AVGADES-----VKEAHREVINSSTEGLLNIDKDIRKILSYIVE 645
Db 654 NTINNIDVINKIGAGDAVLSDYFNIGITDVYQDILDYVNADLKIQNKYKDVDDIIEVREAK 713
QY 646 IEDTEGLK-----EVINDRYDMLNSSLRQDKGKTFFIDFKKYNDKLPY-----ISNENYK- 695
Db 714 ISSYEALMRINIGEAVTDDFKALGLTDI-----NDGLLLYATTLQNKNYKT 760
QY 696 -----VNVY-AVTKENTIIINPSENGDSTNGI 721
Db 761 ADEVIARVQAQIEIYRALMQIN--LGKATTADVNTLGI 796

Search completed: May 3, 2004, 19:41:23
Job time : 17.0855 secs

```

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 8.2539 Seconds  
(without alignments)  
4636.784 Million cell updates/sec

Title: US-09-848-909A-13

Perfect score: 3772

Sequence: 1 EVKQENLLNESSSSGGL.....TSTNGIKILIFSKKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3750	99.4	764	1	PAG_BACAN
2	235.5	6.2	204	1	YBI_BACAN
3	184.5	4.9	1630	1	MSPI_PLAPK
4	184.5	4.9	1639	1	MSPI_PLAPF
5	169	4.5	1658	1	YM67_YEAST
6	168	4.5	1000	1	S155_YEAST
7	166	4.4	1276	1	BXD_CLOBO
8	165	4.4	1803	1	YUL3_YEAST
9	164	4.3	1957	1	SFOF_SCHPO
10	162	4.3	1385	1	FAT1_SCHPO
11	161.5	4.3	1230	1	SMC3_YEAST
12	160.5	4.3	979	1	P115_MYCHR
13	160.5	4.3	1487	1	MSJ3_YEAST
14	160.5	4.3	2334	1	WAPA_BACSU
15	158.5	4.2	1167	1	CAGA_HELPJ
16	158.5	4.2	1233	1	Y116_YEAST
17	158.5	4.2	2116	1	MYS2_DICDI
18	158.5	4.2	2869	1	RBP1_PLAVB
19	157.5	4.2	1024	1	RIP3_MOUSE
20	157.5	4.2	2867	1	RB22_PLAVB
21	156	4.1	918	1	YMTB_CABEL
22	155.5	4.1	1208	1	PCP1_SCHPO
23	155	4.1	1037	1	KC4A_YEAST
24	155	4.1	1790	1	USO1_YEAST
25	155	4.1	1928	1	MYS1_YEAST
26	154.5	4.1	1460	1	N159_YEAST
27	153	4.1	1288	1	VACA_HELPJ
28	152.5	4.0	1046	1	SBCCLACLA
29	152	4.0	1029	1	RIP3_PAT
30	151.5	4.0	1271	1	Y338_MYCGE
31	151	4.0	1250	1	BXE_CLOBO
32	150.5	4.0	1225	1	Y309_MYCGE
33	150.5	4.0	1358	1	SIR4_YEAST

34	150	4.0	2696	1	NSD1_HUMAN	Q96173 homo sapien
35	149.5	4.0	678	1	YNC7_YEAST	P53968 saccharomyc
36	149	4.0	1420	1	SRB9_YEAST	P38931 saccharomyc
37	148	3.9	1577	1	HLVA_PROMI	P16466 proteus mir
38	147.5	3.9	1014	1	HEX2_YEAST	Q00816 saccharomyc
39	147.5	3.9	1091	1	CIC2_PAT	P54290 rattus norv
40	147.5	3.9	1290	1	BXB_CLOBO	P10844 clostridium
41	147.5	3.9	1916	1	RIF1_YEAST	P29539 saccharomyc
42	146.5	3.9	730	1	GLN3_YEAST	P18494 saccharomyc
43	146.5	3.9	2104	1	MYS3_SCHPO	O14157 schizosacch
44	146	3.9	719	1	YM41_YEAST	Q03213 saccharomyc
45	146	3.9	1183	1	CNA_STAAU	Q53654 staphylococ

ALIGNMENTS

RESULT 1

PAG\_BACAN

STANDARD; PRT; 764 AA.

AC P13423; Q9F5R7; Q9KH69; Q9ROU2;

DT 01-JAN-1990 (Rel. 13, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins

DE translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].

GN PAG\_A OR PAG\_OR PXO1-110.

OS Bacillus anthracis.

OG Plasmid pXO1.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI\_TaxID=1392;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89172073; PubMed=3148491;

RA Welkos S.L.; Lowe J.R.; Eden-McCutchan F.; Vodkin M.; Leppla S.H.;

RA Schmidt J.J.;

RT "Sequence and analysis of the DNA encoding protective antigen of

RT Bacillus anthracis.";

RL Gene 69:287-300(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=28, 33, BA1024, and BA1035;

RX MEDLINE=99214082; PubMed=10197996;

RA Price L.B.; Hugh-Jones M.; Jackson P.J.; Keim P.;

RT "Genetic diversity in the protective antigen gene of Bacillus

RT anthracis.";

RL J. Bacteriol. 181:2358-2362(1999).

RN [3]

RP SEQUENCE FROM N.A.

RX STRAIN=V770-NP1-R / ATCC 14185;

RX MEDLINE=20359347; PubMed=1089854;

RA Cohen S.; Mendelson I.; Altboum Z.; Kobiler D.; Elhanany E.; Bino T.;

RA Leitner M.; Inbar I.; Rosenberg H.; Gozes Y.; Barak R.; Fisher M.;

RA Kronman C.; Velan B.; Shafferman A.;

RT "Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus

RT anthracis spore vaccines protect against anthrax.";

RL Infect. Immun. 68:4549-4558(2000).

RN [4]

RP SEQUENCE FROM N.A.

RX STRAIN=Sterne;

RX MEDLINE=99445483; PubMed=10515943;

RA Okinaka R.T.; Cloud K.; Hampton O.; Hoffmaster A.R.; Hill K.K.;

RA Keim P.; Kohler T.M.; Lamke G.; Kumano S.; Mahillon J.; Manter D.;

RA Martinez Y.; Ricke D.; Svensson R.; Jackson P.J.;

RT "Sequence and organization of pXO1, the large Bacillus anthracis

RT plasmid harboring the Anthrax toxin genes.";

RL J. Bacteriol. 181:6509-6515(1999).

RN [5]

RP	DOMAINS.					
RX	MEDLINE=91332080; PubMed=1651334;					
RA	Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;					
RT	"The carboxyl-terminal end of protective antigen is required for					
RT	receptor binding and anthrax toxin activity."					



RL J. Biol. Chem. 266:15493-15497(1991).  
 RN [6]  
 RP CHARACTERIZATION.  
 RX STRAIN=Stierne;  
 RA MEDLINE=94327640; PubMed=80511159;  
 RT Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;  
 "Anthrax protective antigen forms oligomers during intoxication of  
 mammalian cells".  
 RL J. Biol. Chem. 269:20607-20612(1994).  
 RN [7]  
 RP CHARACTERIZATION.  
 RX MEDLINE=21129592; PubMed=11207581;  
 RA Beauregard K.E., Collier R.J., Swanson J.A.;  
 "Proteolytic activation of receptor-bound anthrax protective antigen  
 on macrophages promotes its internalization".  
 RL Cell. Microbiol. 2:251-258(2000).  
 RN [8]  
 RP TOXIN REGULATION.  
 RX STRAIN=Weybridge;  
 RA MEDLINE=94131936; PubMed=8300513;  
 RA Koehler T.M., Dai Z., Kaufman-Yarbray M.;  
 "Regulation of the Bacillus anthracis protective antigen gene: CO2 and  
 a trans-acting element activate transcription from one of two  
 promoters".  
 RL J. Bacteriol. 176:586-595(1994).  
 RN [9]  
 RP MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.  
 RX STRAIN=Stierne;  
 RA MEDLINE=95050722; PubMed=7961869;  
 RA Singh Y., Klimpel K.R., Arora N., Sharma M., Leppla S.H.;  
 "The chymotrypsin-sensitive site, FFD315, in anthrax toxin protective  
 antigen is required for translocation of lethal factor".  
 RL J. Biol. Chem. 269:29039-29046(1994).  
 RN [10]  
 RP MUTAGENESIS OF DOMAIN 4 LOOPS.  
 RX STRAIN=Stierne;  
 RA MEDLINE=99185012; PubMed=10085028;  
 RA Varughese M., Teixeira A.V., Liu S., Leppla S.H.;  
 "Identification of a receptor-binding region within domain 4 of the  
 protective antigen component of anthrax toxin".  
 RL Infect. Immun. 67:1860-1865(1999).  
 RN [11]  
 RP MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.  
 RX STRAIN=Stierne;  
 RA MEDLINE=21092804; PubMed=11178978;  
 RA Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;  
 "Trp 346 and Leu 352 residues in protective antigen are required for  
 the expression of anthrax lethal toxin activity".  
 RL Biochem. Biophys. Res. Commun. 281:186-192(2001).  
 RN [12]  
 RP MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.  
 RX STRAIN=Stierne;  
 RA MEDLINE=21438996; PubMed=11554763;  
 RA Anuja N., Kumar P., Bhatnagar R.;  
 "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are  
 required for oligomerization of anthrax protective antigen".  
 RL Biochem. Biophys. Res. Commun. 287:542-549(2001).  
 RN [13]  
 RP MUTAGENESIS OF PRO-289.  
 RX STRAIN=Stierne;  
 RA MEDLINE=21255689; PubMed=11356563;  
 RA Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;  
 "Role of residues constituting the 2beta1 strand of domain II in the  
 biological activity of anthrax protective antigen".  
 RL FEMS Microbiol. Lett. 199:27-31(2001).  
 RN [14]  
 RP MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.  
 RX MEDLINE=2122576; PubMed=11222612;  
 RA Mogridge J., Mouruz M., Collier R.J.;  
 "Involvement of domain 3 in oligomerization by the protective antigen  
 moiety of anthrax toxin".  
 RL J. Bacteriol. 183:2111-2116(2001).  
 RN [15]

RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.  
 RX MEDLINE=21269403; PubMed=11113126;  
 RA Sellman B.R., Nassi S., Collier R.J.;  
 "Point mutations in anthrax protective antigen that block  
 translocation".  
 RL J. Biol. Chem. 276:8371-8376(2001).  
 RN [16]  
 RP MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;  
 ILE-239; TRP-255 AND PHE-265.  
 RX STRAIN=Stierne;  
 RA MEDLINE=22112896; PubMed=12117959;  
 RA Chauhan V., Bhatnagar R.;  
 "Identification of amino acid residues of anthrax protective antigen  
 involved in binding with lethal factor".  
 RL Infect. Immun. 70:4477-4484(2002).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ÅNGSTROMS).  
 RX MEDLINE=97192099; PubMed=9039918;  
 RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;  
 "Crystal structure of the anthrax toxin protective antigen".  
 RL Nature 385:833-838(1997).  
 RN [18]  
 RP REVIEW.  
 RX MEDLINE=21428689; PubMed=11544370;  
 RA Mock M., Fouet A.;  
 "Anthrax".  
 RL Annu. Rev. Microbiol. 55:647-671(2001).  
 CC -I- FUNCTION: One of the three proteins composing the anthrax toxin,  
 the agent which infects many mammalian species and that may cause  
 death. PA binds to a receptor (APR) in sensitive eukaryotic  
 cells, thereby facilitating the translocation of the enzymatic  
 toxin components, edema factor and lethal factor, across the  
 target cell membrane. PA associated with LF causes death when  
 injected, PA associated with EF produces edema. PA induces  
 immunity to infection with anthrax.  
 CC -I- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a  
 protective antigen (PA), a lethal factor (LF) and an edema factor  
 (EF). None of these is toxic by itself. PA+LF forms the lethal  
 toxin (lethx); PA+EF forms the edema toxin (edtx). PA-63 forms  
 heptamers and this oligomerization is required for LF or EF  
 binding. Once activated, at low pH, the heptamer undergoes  
 conformational changes and converts from prepore to pore inserted  
 in the membrane, forming cation-selective channels.  
 CC -I- SUBCELLULAR LOCATION: Secreted.  
 CC -I- DOMAIN: The molecule is folded into four functional domains. Each  
 domain is required for a particular step in the toxicity process.  
 CC Domain 1 contains two calcium ions and the proteolytic activation  
 site. Cleavage of the PA monomer releases the subdomain 1a, which  
 is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is  
 part of the remaining 63-kDa fragment (PA63) and contains the  
 binding sites for LF and EF. Domain 2 is a beta-barrel core  
 containing a large flexible loop that has been implicated in  
 membrane insertion and pore formation. There is a chymotrypsin  
 cleavage site in this loop that is required for toxicity. Domain 3  
 has a hydrophobic patch thought to be involved in protein-protein  
 interactions. Domain 4 appears to be a separate domain and shows  
 limited contact with the other three domains: it would swing out  
 of the way during membrane insertion. It is required for binding  
 to the receptor; the small loop is involved in receptor  
 recognition.  
 CC -I- PTM: Proteolytic activation by furin or a furin-like protease  
 cleaves the protein in two parts, PA-20 and PA-63; the latter is  
 the mature protein. The cleavage occurs at the cell surface and  
 probably in the serum of infected animals as well; both native and  
 cleaved PA are able to bind to the cell receptor. The release of  
 PA20 from the remaining receptor-bound PA63 exposes the binding  
 site for EF and LF, and promotes oligomerization and  
 internalization of the protein.  
 CC -I- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were  
 performed that showed that the residues present in the small loop  
 of domain 4, and not the ones in the large loop, are involved in  
 receptor recognition.  
 CC -I- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

Query Match 99.4%; Score 3750; DB 1; Length 764;  
Best Local Similarity 99.5%; Pred. No. 1.4e-179;  
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQNRLLNSESSESSQGLLYGFFDLNFPQPMVVTSTTGDLSIPSELENIPSENYF 60  
DB 30 EVKQNRLLNSESSESSQGLLYGFFDLNFPQPMVVTSTTGDLSIPSELENIPSENYF 89  
QY 61 QSAIWSGFIKVKSDYEYFATSAADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 90 QSAIWSGFIKVKSDYEYFATSAADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 149  
QY 121 QRENTEKGLDPLKWTDSQNKKEYISSDNLQPLKQKSSNRKRSSTAGPTVPDRDN 180  
DB 150 QRENTEKGLDPLKWTDSQNKKEYISSDNLQPLKQKSSNRKRSSTAGPTVPDRDN 209  
QY 181 DQIPSPLEVEGYTVKVKRTPSLSPWISNHEKGLTKYKSPKSWSTASPYSDFEKVT 240  
DB 210 DQIPSPLEVEGYTVKVKRTPSLSPWISNHEKGLTKYKSPKSWSTASPYSDFEKVT 269  
QY 241 GRIDKNVPEARHPIVAAPYPIVHVDENIILSKNEDQSTQNTDSTRISKNTSRTHT 300  
DB 270 GRIDKNVPEARHPIVAAPYPIVHVDENIILSKNEDQSTQNTDSTRISKNTSRTHT 329  
QY 301 SEVHGNAEVAHFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 330 SEVHGNAEVAHFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 389  
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
DB 390 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPSKNLAPIA 449  
QY 421 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480  
DB 450 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 509  
QY 481 LPQIQETARIIFNGKDLNLVRRITAAVNPSPLETTKPDMTLKALKIAGFNPENGL 540  
DB 510 LPQIQETARIIFNGKDLNLVRRITAAVNPSPLETTKPDMTLKALKIAGFNPENGL 569  
QY 541 QYQKQDITDFDNFDQOSTQNKQNLAEINATNYITVLDKIKLNAQNMILIRDKFHYDR 600  
DB 570 QYQKQDITDFDNFDQOSTQNKQNLAEINATNYITVLDKIKLNAQNMILIRDKFHYDR 629  
QY 601 NNIAVGADESVEKAEHREVSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660  
DB 630 NNIAVGADESVEKAEHREVSSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 689  
QY 661 DMLNISLSRQDGKTFIDFKKYNKPLYSNPNKYVNVYAVTKENTIIINPSENGDTSTNG 720  
DB 690 DMLNISLSRQDGKTFIDFKKYNKPLYSNPNKYVNVYAVTKENTIIINPSENGDTSTNG 749  
QY 721 IKKILFSSKGYEIG 735  
DB 750 IKKILFSSKGYEIG 764

RESULT 2  
YFBI\_BACAN  
ID YFBI\_BACAN STANDARD; PRT; 204 AA.  
AC P13422; Q9X377;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein pXOI-111.  
GN pXOI-111.  
OS Bacillus anthracis.  
OG Plasmid pXOI.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.

Query Match 6.2%; Score 235.5; DB 1; Length 204;  
Best Local Similarity 34.8%; Pred. No. 1.9e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNLIRDKRPHYDRNNIAVGADESVEKAEHREVSSTEGLLNIDKIRKILSGYVIEI 646  
DB 1 MNLIRDRP-YHYDNGNIVGVDDSYLKNAYKQILNWSGVSINLDEDVQALSGYMLQI 59  
QY 647 EDTE-----GLKEVINDRYDMLNISLSRQDGKTFIDFKKYNKPLYSNPN 693  
DB 60 KPSNHLTNSPVITLAKGDSGVGELYRVL-----DGTGFLDFNKFDEWWSLV-DPG 112  
QY 694 YKVNYYAVTKEN-TIINPSENGDTSTNGIKKILFSSKGYEIG 734  
DB 113 DDVVVYAVTKEDFNAVTRDENGNI-NKLKNTLVLSGKIKEI 153

RESULT 3  
MSPI\_PLAFK  
ID MSPI\_PLAFK STANDARD; PRT; 1630 AA.  
AC P04932;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-PEB-1996 (Rel. 33, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)  
DE (PMWSA) (P190).  
GN MSP-1.  
OS Plasmodium falciparum (isolate K1 / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86136024; PubMed=3004972;  
RA Mackay M., Goman N., Bore N., Hyde J.E., Scaife J., Certa U.,  
RA Stunnenberg H., Bujard H.;

EMBL; M22589; AAA22636.1; -;  
DR EMBL; AF065404; AAD32415.1; -;  
DR PIR; G59104; G59104.  
DR PIR; I39933; I39933.  
DR HSP; P13423; IACC.  
KW Hypothetical protein; Plasmid; Transmembrane.  
FT TRANSEM 162 182 POTENTIAL.  
FT CONFLICT 93 93 T->A (IN REF. 1).  
FT CONFLICT 184 204 KSCNCLIIYEVESQLMNSVY -> NHVIVYLSM (IN REF. 1).  
SQ SEQUENCE 204 AA; 23029 MW; E1657B23AE4273FD CRC64;

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

"Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merozoites: studies at the genetic level."; EMBO J. 4:3823-3829 (1985).

[2]

REVIEWS, SEQUENCE FROM N.A.

Pan W., Tolle R., Bujard H.,  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).

CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

EMBL: X03371; CAA27070.1; -  
DR InterPro: IPR006209; EGF-like.  
DR Pfam: PF00008; EGF; 1.  
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.  
FT DOMAIN 67 84 TRIPETIDE SG(TP) REPEAT.  
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.  
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1630 AA; 187289 MW; ABBDEC3CE0A46322 CRC64;

Query Match 4.9%; Score 184.5; DB 1; Length 1630;  
Best Local Similarity 20.4%; Pred. No. 0.11;  
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

QY 5 ENRL-LNESESSQGLGYFSDLNFCQAPMVVTS\*TGDLSTPSSSELENIPSENQVQSA 63  
DB 918 ENLSLGNKNVYQELIGQSS-NP-----YKILKDSDFYNE 956  
QY 64 IWSGFIVKXSDYFTFATSADNHVTWMDQDVINKASNS-----NKIRLEKGRLY- 114  
DB 957 SFTNFVKSADD-----INSLNDESKKLEEDINKLKTQLSFDLYNKYKLLERLFD 1011  
QY 115 -----QIKIQYQRENTKGLDFKLYWTSQNKKEVSSDNLQLPELKQSSNSRKES 168  
DB 1012 KKTVGVYKQWIKLTKLQLESK-----NSLNPKGVL--QNFVFNKKKEAIEAETN 1067  
QY 169 TSAGTVPDRDNDGIPDSLEVEGYVDVQNKRTFLSPWISNIHKKGLTKY-----KSSP-- 223  
DB 1068 T-----LENTKILLKH-----KGLVKVYNGESSPLK 1094  
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPAR-----HPVAVYPIVHVD 265  
DB 1095 TLSEESITQEDNYSLENFKVLSKLEGLKNDLNLEKKKLSYSSGLHLLHIA-----E 1147  
QY 266 MENILSKNEQDSQNTDSETRTISKNTSTSRHTSEVHGNAEYHASFDFDGGSVAGFS 325  
DB 1148 LKEVIKKN-----YTCNSPSENNT-----DVNNALSEYKFLPEGTDVAITWS 1191

QY 326 NSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPFYNVLPPT---- 381  
DB 1192 EGSOTLSESQPKPASTHVGAE--NTITTSQ-NVDEVDVVIIVIFGESEEDYDDL 1247  
QY 382 -SLVIGKQOTLATIKADENQLSOLLAPNNYPSKNLAPIA---LNAQKDASTPTTMNVN 437  
DB 1248 GQVVTGEAVTPSVI---DNILSKI---ENEVEVLYLKLGLAGYRSLKQLENVNVTFNVN 1301  
QY 438 -----QFLELEKTKQLRLDQV-YGNIAF-----YNFENGRRVYRVDTSQNMSEVL 481  
DB 1302 VKDILNSRNKNEPKV-LESDLPYKDLTSSNVVVDPKYFLNKEKDKFLSYNYIK 1360  
QY 482 PQIQETARIIFNGKDLNLVERRIAAVNPSDF-----ETTKPDM-TLKEALKIAGF 533  
DB 1361 DSID-----TDINFA-----NDVLGYKILSEKYSKSLDSIKYINDKQGE 1401  
QY 534 NEPN-----GNLOVQGDITE-----DFNFQDQTSQNTKQNLABELNATNYT 576  
DB 1402 NEKYLFLNNIETLYKTVNDKIDLFIHLEAKVLNVYEK---SNVEVKIKELN--YLKT 1456  
QY 577 VLDKIKLNAMN--ILIRKRFHYDRNN-----IAGV-ADESVVKEAHEVINSSTEGLL 628  
DB 1457 IQDKLADFKNNFVGIADLSTDYNNHLLTKFLSTGMVFENLAKTVLSNLDGNLQGM 1515  
QY 629 LNIDKD--IRKILSGYVIEIDTEGLKEVINDRYDMLNLSLRQDQKTFIDPKYNDKLP 686  
DB 1516 LNISSQCVKK-----QCPQNSGGCFRHLDE-----REECKCILNYKQEGDKC- 1557  
QY 687 LYISNPYKVNVAVTKENTINPSNG-----DTSTNGIKKI 724  
DB 1558 --VENFNPCTN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1590

RESULT 4  
MSPI\_PLAFW  
ID MSPI\_PLAFW STANDARD; PRT; 1639 AA.  
AC P04933;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens) (PMWSA) (P195).  
GN MSP-1.  
OS Plasmodium falciparum (isolate Wellcome).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
RN [1] NCBI\_TaxID=5948;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86014355; PubMed=2995820;  
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V., Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.I., Freeman R.R.;  
RT "Primary structure of the precursor to the three major surface antigens of Plasmodium falciparum merozoites.";  
RL Nature 317:270-273 (1985).  
RN [2]  
RP REVISIONS.  
RA Holder A.A.;  
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).  
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merozoites. The maturation take place during schizont.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; X02919; CAA26676.1; --  
DR PIR; A24594; A24594.  
DR PIR; S05603; S05603.  
DR PDB; 1CE3; 28-MAY-99.  
DR InterPro; IPR006209; EGF\_like.  
DR Pfam; PF00008; EGF; 1.  
DR Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor; 3D-structure.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1639 MEROZOITE SURFACE PROTEIN 1.  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 768 768 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 783 783 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 844 844 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 920 920 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F66 CRC64;  
Query Match 4.8%; Score 184.5; DB 1; Length 1639;  
Best Local Similarity 20.5%; Pred. No. 0.11; Mismatches 256; Indels 265; Gaps 47;  
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;  
QY 5 ENRI-LNESSSSGLLGYYFSDINFAQPMVVTSTTGDLSIPSSSELENIPSNQYFQSA 63  
DB ENILSLGKNKIYQELIGKQSE-NF-----YEKILKSDTFYNE 965  
QY 64 IWSGFIKVKSDYFATSDNHTVMVDQEVINKASNS-----NKIRLEKGRLY- 114  
DB SFTNFVSKADD-----INSLNDESKKXUEBINKLKTQLSLFDLYNKYKILKLERLD 1020  
QY 115 -----QIKIQYQRENPTKGLDPLKYWDSQNKKEVISSDNLQPLKOKSSNSRKRKS 168  
DB KKTVGKYKMQIKLLKEQLSKL-NSLNPKFVL-QNFSVFNNKKKEAIEAETN 1076  
QY 169 TSAGPTVPDRNDGIPDSLEVEGVTVVKNKRTFLSPWISNIHKKGLTKY--KSP-- 223  
DB T-----LENTKIILKHY-----KGLVYKNGESSPLK 1103  
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPEAR-----HPLVAAPYIHVVD 265  
DB TLSESIQTDENYASLENFKVLSKLEGLKDNLNLEKKLSYLSGLHLIA-----E 1156  
QY 266 MENIILSKNEPQSTQNTDSETRTISKNTSTSRTHSEVHGNAEVHASFIDIGSVSAGFS 325  
DB 1157 LKEVIKKNK---YTGNSPSENNT-----DVNNALSYKKFLPEGTDTVATWS 1200  
QY 326 NSNSTVALDHSLSLAGERTWAEITGLTADTARLNANIRVNTGTAPIYNVLPFT--- 381  
DB 1201 EGGSTLQSQSPKPFAPSTHVAES---NTIITSQ-NVDEVDVLIIVIFGESEDDDL 1256  
QY 382 -SLVLGRKQTLATIKADENQLSILAPNNYPSKXNLAPIA---LNAQKQASSTPITNNYN 437  
DB 1257 GOVVTGEAVTPSVI---DNILSKI---ENEYEVLYLKPLAGVYRSLKKQLENNVMTFNVN 1310  
QY 438 -----QFELEKTKQRLDQV-YGNIAI-----YFNGRVRVDTGSNWSEVL 481  
DB 1311 VKDILNSRFNKGKFNKFN-LESDDLIPYKDLTSSNVVVKDPYKFNKKEKDFLSYNIYK 1369  
QY 482 PQIQETARIIFNGKDLNVERRIAANVNSDPL-----ETTKPDM-TLKEALKIAPGF 533  
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSDDLDSIKKYINDKQGE 1410  
QY 534 NFPN-----ONLYQOKDITE-----PFFNFQDQTSQNIKNQLAEILNATNIYT 576  
DB 1411 NEKYLFFLNIIETLYKTNDRKIDLFVHLEAKVLNITYEK-----SNVEVKIKELN---YLDK 1465

QY 577 VLDKIKLNKAKN--ILIRDKRFHYDRNN-----IANG-ADESVVKEAHREVINSGTEGLL 628  
DB 1466 IODKLADPKONNFVGIADLSTDYNNHLLTKFLSTGMVFENLAKTFLVLSNLDGNLOG-M 1524  
QY 629 LNIKD--IRKILSGYIEIEDTGLKEVINDRYDMLNLISSLRDQGTDFDKYNDKLP 686  
DB 1525 LNISSHQCVKK-----QCQNSGCFRHLDE-----RECKLLNYKQEGDKC- 1566  
QY 687 LYISNPYKVNYYAVTKNTIINPSENG-----DTSTNGIKKI 724  
DB 1567 --VENPNTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1599  
RESULT 5  
ID Y67 YEAST STANDARD; PRT: 1658 AA.  
AC Q03661; Q04988;  
DT 01-NOV-1997 (Rel. 35; Created)  
DT 01-NOV-1997 (Rel. 35; Last sequence update)  
DT 10-OCT-2003 (Rel. 42; Last annotation update)  
DE Hypothetical 187.1 kDa protein in GUA1-ERG8 intergenic region.  
GN YMR219W OR YMR261.13 OR YMR955.01.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / AS972;  
RX MEDLINE=97313268; PubMed=9169872;  
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,  
RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,  
RA Jagels P., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,  
RA Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome  
XIII.";  
RT Nature 387:90-93(1997).  
RL -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z49809; CAA89934.1; --  
DR EMBL; Z49339; CAA90190.1; --  
DR PIR; S55101; S55101.  
DR Germonline; 142894; --  
DR SGD; S0004832; ESCI.  
DR GO; GO:0005634; C:nucleus; IDA.  
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.  
KW Hypothetical protein.  
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;  
Query Match 4.5%; Score 169; DB 1; Length 1658;  
Best Local Similarity 18.6%; Pred. No. 0.66;  
Matches 159; Conservative 140; Mismatches 274; Indels 284; Gaps 41;  
QY 1 EVKQENRLNSES-----SSQGLLGYFSDINFAQPMVVTSTTGDLSIPSSSELENIPSEN 57  
DB 290 EMELEDDIDVESDAEKDESQAGIEHS-VDFSKYMPRTDNTKIPVIEKYESDEHKVHQ 348  
QY 58 QYQPSAHS-GFTKVKKSE-----YTPATSDN-----HVTMWVDDQEVINKASNSK 105  
DB 349 RYSGDGFDFGYSINIVDDDESEDEBSQAESYANAEVYHNEHLEDDKELIEDIESDS 408  
QY 106 IRLEKRLVQIKIQYQRENPTKGLDFKL-----YWTDSQKKEV 145  
DB 409 -----ESQSAQSESGESDDPFYKMKNEKSTSEETNTSESDQGFADYTKNV 459  
QY 146 ISSDNLQLPEL-----KQSSNSRKRKRSAGTVPDRNDGIPDSLEVE 190

```

Db 460 EQQENDEPEKDDIIRSLDKNFHGNKNSYSENVLENEDFPAIVERNQ-IND-----VE 515
Qy 191 GYTDVQKQKRTFL---SPWISNIIEH--KKGITKTKSS-----PKWSTASD----- 231
Db 516 GYDVTGKSEDLHEHSP--DNLVDLAARAMLQFQQGRNSNCPOKEQVSESYLGHSHNGS 573
Qy 232 -----PYSDFEKTGRIDIKQVSPPEARHPLVAAYPIVHVDMENIILSKNEDQ 277
Db 574 NLSGSLDESBEQIPLKDF--TGNNNNLKT-DGDLSSS--VEIEVEKV---SEKK 622
Qy 278 STQNDSTSTRISKNTSTSRTHITSEVHGNAEVAHSPFDIGGSVAGFSNNSSTVALDHS 337
Db 623 LGSGTEKELVPLSTDTTN-----NSSLGDEDSIYISLDD- 657
Qy 338 LSLAGERTWAEITMGLNTADTARLANIRYVNTGAPINVLPTSLVLGKQNTIATKAD 397
Db 658 -----ADAISENLTD-----VPLMEIKTKPKYEVVISESV-----YSSISYE 694
Qy 398 ENQUS-----QILAP-----NNYIPSKN-----LAPIALNAOKDAS----- 428
Db 695 DNTVAMPQQVEYTSFPMNDPFFNSLNDYKXKHLKSTLAALAPAF--TKDAEFVEAGV 752
Qy 429 ---STPTTMVQNQLEKTKQLRLDQVYGNIAVTNPFNGRVRVDTGSGNWSEVLPOI 484
Db 753 TKSLTSTSGHNIHFTSKETKQVS--DLDESTENVTENNTG-----DENKQSKNFGV 807
Qy 485 QETTARIIFNGKDLNLVERRIAANVPSPLETTKPDMLKALKIAPFNPNGNLQYQG 544
Db 808 ANSTDKSTEDNTD---EKYFSAINYTN-----VTGDSCEDIETASNVBE--NLRYCE 856
Qy 545 KDITFDFNF--DOOTSON---IKQLA-----ELNATNIYTVLDKIKLNAKNMILRD 593
Db 857 KDMNEAESGGEVCKQNDGSKTQISFSTSPDNFQESNDNTEFSSTK----- 905
Qy 594 KRFHYDRNIAVGADSEVYKGAHR--EVIN-----SSTEGLLNID 632
Db 906 ---YKVRNSDLEDDSESLKELTKAEVVDKLEDESEDSYEQDYADPEPGNDEGSNIV 961
Qy 633 KDRIKILSGYIVEIEDTGLKEVINDRYDMLNLISSLRQDGKTFIDFKYNDKLPYISNP 692
Db 962 KGTKK-----DTGIVEPENEKYN-----KVHEETLFFBANV 993
Qy 693 NYKANVYAVTKENTIIN 709
Db 994 SSSVANVQNKMDTVDIN 1010

RESULT 6
S155_YEAST
ID S155_YEAST STANDARD; PRT; 1000 AA.
AC P43612;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S154-associated protein SAPI55.
GN SAPI55 OR YFR040W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96220458; PubMed=8649382;
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Arndt K.T.;
RT "The SAP, a new family of proteins, associate and function positively
RT with the S154 phosphatase."
RL Mol. Cell. Biol. 16:2744-2755(1996).
RN [2]
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;

```

```

RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=96287652; PubMed=8686379;
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT "Analysis of a 36.2 Kb DNA sequence including the right telomere of
RT chromosome VI from Saccharomyces cerevisiae.";
RL Yeast 12:149-167(1996).
CC -!- FUNCTION: Associates with the S154 phosphatase in a cell cycle
CC dependent manner. May be directly or indirectly involved in
CC S154-dependent functions in budding and in normal G1 cyclin
CC expression.
CC -!- PTM: Hyperphosphorylated in the absence of S154.
CC -!- SIMILARITY: Belongs to the SAPS family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U50560; AAC49303.1; -.
CC EMBL; D50617; BAA09279.1; -.
CC GenBank; 140194; -.
CC SCD; S0001936; SAPI55.
CC InterPro; IPR007587; SAPS.
CC Pfam; PF04499; SAPS; 1.
CC Phosphorylation; Cell cycle.
KW CONFLICT 663 663
FT CONFLICT 674 699
FT REF. 2 AND 3).
FT REF. 2 AND 3).
FT CONFLICT 814 815
FT CONFLICT 818 822
FT CONFLICT 818 822
FT SEQUENCE 1000 AA; 114880 MW; AD4FBF08C7C586D5 CRC64;
Query Match 4.5%; Score 168; DB 1; Length 1000;
Best Local Similarity 19.5%; Pred. No. 0.37;
Matches 173; Conservative 128; Mismatches 323; Indels 262; Gaps 42;
Qy 1 EVKQENRLNSESSESSQGLGYFS--DL-----NFQAPMVVTSSTGD-----LSIPS 47
Db 90 EKEEERSMNSSESSTFSSTGSKTDLDEEDIGNATAPMVVTKNLDNSPIERMLVET 149
Qy 48 SELENIPSENOYFQSAIWSGFIKVKKSDYTTATSGADNHVTWVDD-QEVINKASNSKI 106
Db 150 ELLNELSRQNKTLDFICGFFGFDKTKNKV-----NNMEYLVQLMCEISKITATTV 203
Qy 107 RLEKGRLYQIKYQRENTTEKGLDFKLVWTSQNKKEVSSDNQLQPLPKOKSSNRK 166
Db 204 DLNN-----LIDYQEQQLD-----DSSQEDVYVESDTEQEKEEDNNNNKK 247
Qy 167 R-----STSAGPTVPDRDNDGIPDSLEVEGYTVDVKNKRTFLSP-----WISNIHEKKG 215
Db 248 RRVKSGSSFGNDDINNDDDDDDANEDDESAYL-----TKATIISEIFSLDIWLIS-----BS 299
Qy 216 LTKYKSPKX-NTASDPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKN 274
Db 300 LVKNQSVLKNKWSIINQ-----NFNSE-NSPLV---PIFLKQNLILLTR- 341
Qy 275 EDQ-----STQNTDSET-----RTISKNT 293
Db 342 ODQYLNFIPTERSFVDDMLKHVDISLLMDFFLKILISTDKIESPTGIIELVYQNLISKCL 401
Qy 294 STGRTHITSEVHGNAEVAHSPFDIGGSVAGFSNNSSTVAI-DHSLS--LAGERTWAEITM 350

```

Db 402 SFLANKESP-----ADIQACVGFALKALIASANAPLDDISIGPNSLTRQLASPESTAKLV 457  
 QY 351 GLNTADTARLNANRYVNTGTAPYVNVLPPTSLVLGKNOTLATIKADENQLSOLAPNNY 410  
 Db 458 DI-----MINORGAALNT-----TVSIVI-----ELIRKNSDYDQVNLTTTIXT 498  
 QY 411 YPSKNLAPIALNAOKASSPTITVNNYQFLEKTKOLRLDQVYGNATYVNFGRVYR 470  
 Db 499 HPPSNRPDIYGLYLLRFSNHLSDFFQIILDIENDANIPLHENQLKHEKFKPLGFERFKV- 557  
 QY 471 VDTGSNNSEVLVPOIQETIARIIFNGKDLNLVERIAAVNP--SDPLETTKPDWTL--KEA 526  
 Db 558 -----VELIAELHCSNGLMNSKAERIAARDKVRQLSHLGALNDLSIEEKEQ 610  
 QY 527 LKIAFG-----FNEPENGVLQVGKDI-----TEFDNFPD---QOTSQNIKNQLAE 568  
 Db 611 LKTHSPETRDTHDLKNNNGKIDNNDNDDSDYGEIDESPFIPIYNNKQNIKLTDP 670  
 QY 569 LNATNIYTVLDKIKLNKAKWTLILDKRPHYDRNNIAVGAD-----ESVWKEAREVIN-- 621  
 Db 671 TVGT-----YKSNVMIKD-----CFQNNELFLTHPWNFWHNVIFDIIOQIFNGR 716  
 QY 622 ---SSTEGLLNID--KDIRKILSGYIVIEB---DTEGLKEVINDRYDMLNISS---LR- 669  
 Db 717 MDFSYSNLSFLVSLFNLSYQFMTDIVISDEKGTDSRFSVIRDPNDFDKITITDFILRG 776  
 QY 670 -QDKTFTIDFKYN-----DKLPLYLSN-----PNYKNVY 699  
 Db 777 YQDSYKFFELRKWNLGYMGHVLIAEYVVKFSKLYKVDYISRYSNLQTEWQYISSEVL 836  
 QY 700 AVTK-----ENTINPSENGDSTN 719  
 Db 837 NETRMYSKILGGSYIDGNGNIIPQLPDNTTVLTP--NGDASNN 880

## RESULT 7

BXD CLOBO  
 ID - BXD CLOBO STANDARD; PRT; 1276 AA.  
 AC P19321;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)  
 DE (Bontoxilysin D)  
 GN BODT.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BVD/-3;  
 RX MEDLINE=91016853; PubMed=2216736;  
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,  
 RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;  
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum  
 RL neurotoxin type D.";  
 RL Nucleic Acids Res. 18:5556-5556 (1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CB16;  
 RX MEDLINE=93042276; PubMed=1420572;  
 RA Sunagawa H., Ohyanu T., Watanabe T., Inoue K.;  
 RT "The complete amino acid sequence of the Clostridium botulinum type D  
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding  
 RT phase d-16 phi genome.";  
 RL J. Vet. Med. Sci. 54:905-913 (1992).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RC STRAIN=D-SA, and D-1873;  
 RX MEDLINE=89339741; PubMed=2668193;  
 RA Moriiishi K., Syuto B., Kubo S., Oguma K.;

"Molecular diversity of neurotoxins from Clostridium botulinum type D strains.";  
 RL Infect. Immun. 57:2886-2891 (1989).  
 RN [4]  
 RX IDENTIFICATION OF SUBSTRATE.  
 RA MEDLINE=94230352; PubMed=8175689;  
 RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,  
 RA Roques B., Eykne E.M., Suedhof T.C., Jahn R., Niemann H.;  
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and  
 RT F botulinum neurotoxins and tetanus toxin.";  
 RL J. Biol. Chem. 269:12764-12772 (1994).  
 CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS--LEU-61 BOND OF  
 CC SYNAPTOSOMAL VESICLES AND -2.  
 CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
 CC detected action on small molecule substrates.  
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -!- SUBUNIT: disulfide-linked heterodimer of a light chain (L) and a  
 CC heavy chain (H). The light chain has the pharmacological activity,  
 CC while the N- and C-terminal of the heavy chain mediate channel  
 CC formation and toxin binding, respectively.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- MISCELLANEOUS: There are seven antigenically distinct forms of  
 CC botulinum neurotoxin; Types A, B, C1, D, E, F, and G.  
 CC -!- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D  
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE  
 CC BACTERIOPHAGE.  
 CC -!- SIMILARITY: Belongs to peptidase family M27.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMVL; X54254; CAA38175.1; -  
 CC EMVL; S49407; MAB24244.1; -  
 CC PIR; S14455; S11455.  
 CC HSP; P10845; 3BTA.  
 CC MEROPS; M27.002; -  
 CC InterPro; IPR008985; ConA like lec gl.  
 CC InterPro; IPR002160; Kunitz legume.  
 CC InterPro; IPR006025; Pept M-Zn BS.  
 CC InterPro; IPR000395; Peptidase M27.  
 CC Pfam; PF01742; Peptidase M27; 1.  
 CC PRINTS; PR00760; BONTOLYLISIN.  
 CC PRODOM; PD001963; Bontoxilysin; 1.  
 CC PROSITE; PS00142; ZINC PROTEASE; 1.  
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
 CC CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.  
 CC METAL 443 1276 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.  
 CC ACT\_SITE 229 230 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC METAL 230 233 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC DISULFID 437 450 INTERCHAIN (PROBABLE).  
 CC VARIANT 15 16 ND -> PV (IN STRAIN D-1873).  
 CC VARIANT 17 18 K -> Q (IN STRAIN D-SA).  
 CC VARIANT 452 452 R -> T (IN STRAIN D-SA).  
 CC VARIANT 457 457 R -> F (IN STRAIN D-1873).  
 CC VARIANT 462 462 A -> D (IN STRAIN CB16).  
 CC VARIANT 489 489 K -> N (IN STRAIN CB16).  
 CC VARIANT 644 644 N -> K (IN STRAIN CB16).  
 CC VARIANT 1122 1122 O -> R (IN STRAIN CB16).  
 CC SEQUENCE 1276 AA; 146871 MW; C1EC50F46C8233E2 CRC64;

Query Match 4.4%; Score 166; DB 1; Length 1276;  
Best Local Similarity 20.7%; Pred. No. 0.65;  
Matches 165; Conservative 117; Mismatches 290; Indels 224; Gaps 41;

QY 44 SIPSELENIPENOYFOSAIWSGFIKKYKSDVFTATSADNHTVMVDDQEVINKASNS 103  
DB 310 TIPSSWISNI---DKY--KKIFSEKYNFKDNTGNFVNIDKFNLSYDLTNWSEVVYS 364  
QY 104 NKIRLEK-----GRLYQYKIQVQRENPIEKGLDFKLYWTDSONKKEV 145  
DB 365 SQYVKNRTHYPSRHYLPVFANILDNDNIYTRDGF--NLTKGNFIE----- 409  
QY 146 ISSDNLQL-PEUKOKSSNRKRSAGPTVPDRDNDGIPDSLEVEGYVDVKNKRTFLS 204  
DB 410 NSGQNIERNPALOKLSSSVVDLFTKVLRLTKNSRD-----DSTCIKVKNNRL--- 458  
QY 205 PWISN-----IHKKGLTKYKSSPEKWSSTADSPYS-DFEKVGRIDKNVSPPEARPLV 256  
DB 459 PYVADKDSISQIFENKIITD-ETVQNY---SDKFSLDESILDCQVP--INPEIVDPL- 511  
QY 257 AAYPIVHVDMENILSKNE---DQSTQNTD-----SETRTISKNTS--TSRTHUSEV 303  
DB 512 ----LPVNMENPLNPGBEIVFYDITKYVDVLYNSYYVLESQKLSNNVENTLITTSVEEA 567  
QY 304 HGNAEVHASFF-----DIGGSVSAFGSNSNSTVAIDHSLSLAGERTWAETM----- 350  
DB 568 LGVSNKIYTFPLSLAEKVNKGQVAGLFLNANEVVEFTNIMKDTLTKISDVSVIIPY 627  
QY 351 ---GLNTADTA-RLNANIRYVNTG-----TAPIYVLPPTSLVLGKNQTLATIK 395  
DB 628 IGPALNIGNSALRGNFQNAFAGVAFLEGPEFTIPALGVFTFYSSIQEREKIITIE 687  
QY 396 -----ADENQLSQILAPNNY-----YPSKLNAPIALNAQKDASTPTIM 434  
DB 688 NCLEQRVKWQDSYQWVSNLRSITTFQFNHINYQWYDS-----LSVQADAIAKAKIDL 740  
QY 435 NYNQFLEK-----TQQLRLDTPQVQNI-----ATYFENGVRVDTGSN 476  
DB 741 EYKYSQSKENIKSOVENKNSLDVKISEAMNINKFIRECSVTYLFKN----- 790  
QY 477 WSVLPQIQETTARIIFNGKD--LNL-----VERRIAAVNPDPLETTKPMTL 523  
DB 791 ---MLPKVIDELNKPDLRTKTELINLIDSHNIIIVGEVDRKAKYNES--FENTWP---- 841  
QY 524 KEALKIAPG-FNEPKNLOVQKDIETEPFNFQDQTS-----QNIKNQALNATNI-YTV 577  
DB 842 -----FNIFSYTNLSL---KDIINEYFNSINDSKILLSQNKKNALVDTSGYNAEVRV 891  
QY 578 LDXIKLNAXONILIRKFRHYDRNNIAGVADSVVKEAAREVINGSTEGLLINIDKIRK 637  
DB 892 GDNVQLN---TIYTNDFKLSSGDKIIVLNNNILYSA---IYENSVSFWIKISKDLTN 945  
QY 638 ILSGYIV-BIETEGLEKVINDRYDMLNMISSLRQDGKTFIDFKYNDKPLVY----ISNP 692  
DB 946 SHNEYTIINSIEQNSGWKLIRNG---NIEWILQDVN-----RKYKSLIFDYSLSGHT 996  
QY 693 NYKVNVYAVTKENTII 708  
DB 997 GYTNKWFVFTITNNIM 1012

## RESULT 8

YJL3 YEAST  
ID YJL3 YEAST STANDARD; PRT; 1803 AA.  
AC P47024; P87192;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Transposon Ty4 207.7 kDa hypothetical protein.  
GN Ty4B OR YJL113W OR J0780.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1679;  
RX MEDLINE=97103775; PubMed=8948101;  
RA Czepluch C., Kordes E., Pujoil A., Jauniaux J.-C.;  
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X  
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,  
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,  
RT three remnant delta elements and a Ty4 transposon.";  
RL Yeast 12:1471-1474(1996).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; Z49389; CAA9409.1; -.  
CC SGD; S0003649; YJL113W.  
CC InterPro; IPR001594; Rve.  
CC InterPro; IPR001878; Znf\_CCHC.  
CC Pfam; PF00665; rve; 1.  
CC SMART; SM00343; Znf\_C2HC; 1.  
CC Transposable element; Hypothetical protein.  
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;

Query Match 4.4%; Score 165; DB 1; Length 1803;  
Best Local Similarity 19.5%; Pred. No. 1.2; Indels 270; Gaps 41;  
Matches 162; Conservative 123; Mismatches 274; Indels 270; Gaps 41;

QY 1 EVKQNRLLNESSSQGLLYFYFDLNFQAPMVTSTTGDLSPSSSELENIPSNQVY 60  
DB 1078 QLAKTN---HETSPFKGSGIG---TNVKEFNTNNEISLTKTGTSLPTIKLESINHH--- 1128  
QY 61 QSAIWSGFIKKYKSDVFTATSADNHTVMVDDQEVINKASNKIRLEKGLYQIKCY 120  
DB 1129 -----SNDY-----STNKVE-----KF 1140

QY 121 QREN---PTEKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRSAGTVP 176  
DB 1141 EKENHHPPIEDIVDM-----SDQTMESNCQDGNLKLKLVTKVDTGNTGNTVSPR-- 1193  
QY 177 DRNDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKK-----GLTKY----- 219  
DB 1194 -----LEQNEIASGSPQTVNKSFAFLNKEFSLNKKRKRKRHDKNNSLTSYLEDKKR 1247  
QY 220 -----KSSPEKWSSTADSPYSDFEKTGRIDKNVSPPEARPLVAAVPIVHVDMENILSK 273  
DB 1248 SKQNVKLIIDNNMETVSAPKIRAIYNEAISKNPDLKERHEYKQAY---HKEQLNLDKMK 1304  
QY 274 NEDQSTQNTQSETRTTSKN---TSTSRTHSEVHGNABVHASFFDIGGSVSAFGSNSST 331  
DB 1305 VFVDVVKYSRSE---IPDNLIVPTNTIFTKRRNG---IYKARIVCRGDQSPDTSYVIT 1358  
QY 332 VAIDHS-----LSLAGERT-WAETMGLNTADT-ARLNANIRYVNTGTAPIYVLP-TTSL 383  
DB 1359 ESLNENHIKFLMIANNRMFMKTLIDINHAFLAKLEE-----IYIHPHDDRC 1408  
QY 384 VLGNQTLATIKADENQ-----LSQI-LAPNNYYP-----SKNLAPIALNAQKAS 428  
DB 1409 VVKLKALYGLKQSPKEWNDHLRQYLNIGLKQNSYTPGLYQYOTEDKNLM-IAVYVDDCVI 1467  
QY 429 STPTIMYNOFL-ELEKTKOLR-----LDTDOVYGNIAIYNFENGVRV----- 469  
DB 1468 AASNEQRLDEFINKKFNENFKLTGTGLIDVLDTD-ILGMDLYNKKELGTIDITLKSFIN 1526  
QY 470 RVDTCGNNSEVLPQIQTETARIIFNGKDLNLFERRIAAVNP-SDPLETTKPD-----MTL 523  
DB 1527 RMD--KKYNEELKIRKSSIPHMSTYK-----IDPKKDVLMQSEEFQGVKLK 1573

QY 524 KEALKIAGFNBNGLNLOYGKDIETEDFNFDQOTSONIKNQLAELNATNIYTVLDKIKL 583  
 Db 1574 QQLL-----GELNVV-RHKCRDYIEFAVKVAVLVNYPHERVFMIVKIIQYL-- 1620  
 QY 584 NAKMILIDKRFHYDN-----NIAVGADSVVKEAHE-----VI-----NS 622  
 Db 1621 -----VRYKDIGHYDNCNKKVIAITDASVGEYDAQSRIGVILWGMNIFNVYGNK 1675  
 QY 623 STEGLNLNDKDIRKILSY-----IVEIEDT-----EGLK----- 653  
 Db 1676 STNRCVSSAEALHAIYEYADSETLKVTLKELGEGDNDIWIITDSPAIGLNRSYQQ 1735  
 QY 654 -----EVINDRYDMNLNSSLRQDGKTFI-----DPKKY 681  
 Db 1736 PKXKPTWIKTEIKKIKESIKLKITGKNIGADLLTKPVSASFKR 1784

RESULT 9  
 SPOF SCHPO STANDARD; PRT; 1957 AA.  
 ID SPOF SCHPO  
 AC Q10411; Q9USE9;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Sporulation-specific protein 15.  
 GN SPO15 OR SPAC1F3.06C  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=20107136; PubMed=10639340; Shimoda C.;  
 RA Ikemoto S., Nakamura T., Kubo M., Shimoda C.;  
 RT "S. pombe sporulation-specific coiled-coil protein Spo15p is localized  
 RT to the spindle pole body and essential for its modification";  
 RL J. Cell Sci. 113:545-554(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hanlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Ocell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,  
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe";  
 RL Nature 415:871-880(2002).  
 RN [3]  
 RP SEQUENCE OF 705-871 FROM N.A.  
 RC STRAIN=968 h90;  
 RX MEDLINE=20223868; PubMed=10759889;  
 RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,

RA Hiraoka Y.;  
 RT "Large-scale screening of intracellular protein localization in living  
 RT fission yeast cells by the use of a GFP-fusion genomic DNA library";  
 RL Genes Cells 5:169-190(2000).  
 CC -!- FUNCTION: Has a role in the initiation of spore membrane  
 CC formation.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Spindle pole body.  
 CC -!- SIMILARITY: Belongs to the MPC70 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL; Z70690; CAA94624.1; -;  
 DR EMBL; AB027811; BAA87115.1; -;  
 DR F01; T38077; T38077.  
 DR GeneDB SPombe; SPAC1F3.06c; -;  
 KW Sporulation; Coiled coil.  
 FT DOMAIN 199 785 COILED COIL (POTENTIAL).  
 FT DOMAIN 804 1235 COILED COIL (POTENTIAL).  
 FT DOMAIN 1320 1471 COILED COIL (POTENTIAL).  
 FT DOMAIN 1481 1723 COILED COIL (POTENTIAL).  
 FT DOMAIN 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;  
 SQ SEQUENCE

Query Match 4.3%; Score 164; DB 1; Length 1957;  
 Best Local Similarity 20.4%; Pred. No. 1.5;  
 Matches 177; Conservative 132; Mismatches 328; Indels 232; Gaps 40;

QY 2 VKQENRLINSE-----SSSQGLLYGYFDLNFQ---APMVYTSSTGDLSPSSR 49  
 Db 129 VTQKSNLNLKQVRSKLALEHNGILSLQLSSNKKNTSVTLTSEEDVSFQKK 188  
 QY 50 LENIPSENQYFQSAINS---GFIKV-----KKSDYTFATSDAHNVYMWVDDDEVINKAS 101  
 Db 189 LTNMESNFSKQSEAYDLRSQLLTVTEKLDKEDYKEDVSSIKASLAERQASNKSL 248  
 QY 102 NSNKIRLEK---GELYQIKYQRENTE---KGLDFKLYWTDSQNKKEVISSDNLQPE 155  
 Db 249 RGEQERLEKLLVSNKTVTLROTENSLRAECKTLQEKL-----EKAINESDKLLEE 302  
 QY 156 LKQKSSN-----SRKKRSTSGPTVPDRNDGIPDSLEVEGYTVDVYVNRK----- 200  
 Db 303 LKENVANSYDAIVHKDKLIEDLSRISEFN-----LKSERDTLSIKNEKLEKLLRNTI 356  
 QY 201 -TFLSPMISNIHEKKGLTKYKSS-----PEKWTASDPYSDFEK-----VTGRID----- 244  
 Db 357 GSLKDSRTSNLSQLEENVEKESNRTHISQTLTDAESKLSFQENKSLKGSIDYQNNLS 416  
 QY 245 -----KNVS---PEARHPLVAA-----YPIVHVDN----- 266  
 Db 417 SKDRNVQVSQLEEARSSLAHATGKLAENISERDFQNKIKDFEKLQDLRACLNSSN 476  
 QY 267 ---ENILSKNEDQSTQNTD---SETTISKNTSTS-----RHTSVHGNAEVHASFDD 315  
 Db 477 ELKEKSLIDKKQDELNNLREQIKQKQKVSSESTQSSLSQSLQDRILNEKKKHEVYESQLNE 536  
 QY 316 IGGSVSAGFSNSN-----STVAIDHSLSLAGERTWAETMGLTADTARLANANRYVNTG 370  
 Db 537 LKGLQTEINSEHLSQLSTLAKEKAATVNNELSESK---NSLQTL-CNA----- 585  
 QY 371 TAPIYVLPFTSLVLGNQK---TLATIKADENQLSQILAPNNYPSKNLAPIALNAQKD 426  
 Db 586 ---FOEKLAKSVMLKENEQNFSSLDTSFKKLNEHQLENNHQTITKQL-----KD 634  
 QY 427 ASSTPIITMYNQFLEKTKQLRLDTPQVYGNATYFNENGRVYRVDYTSNWSVLPQIQE 486  
 Db 635 TSS-----KIQQLQERANFQKSKSTLSDENNDLRTK-----LLKDEE 672



```

QY 487 TTARIIFNGKDLNLVERRIAAVNPSPDETTKPDMLTKEALKI-----AFGNEPENGNIQ 541
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 673 SNKSLIKKQSDVDSLEKNI-----QTLKEDLRKSEALRFSKLEAKNLREVIDNL- 722
QY 542 YQKQDIT-EPDFNFDQOTSONIKQALAEINATNIYTVLQIKLNKAKWNLIRDKRPHYDR 600
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 723 -KQKHTLEAQRNDLHSSUSDANKTNAILSELTKSSSDVRLTANVELTQDSK----- 776
QY 601 NNIAGADESVVKEAREVINSSTEGLLNIDKDIR-----KILSGYIVEIE-----DT 649
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 777 -----AMQSFSTSLVNSVQS--ISNLYHELADDDHVNQSQNTLLESESKLKTDC 824
QY 650 EQLKE---VINDRYDMLNTSSLRQDGKTIDFKYNDKPLVINSNPYKNVYATKENT 706
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 825 ENITQQNTLIDNVQKLMKHVNQESKV-SELKEVNGKLSLDLKNRLSSNLV-AISDNDQ 882
QY 707 IINP-----SENGDT-----SNGIKKI 724
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 883 ILTQLAELSKNYDSLQESQSLNGLKSL 911

RESULT 10
FAT1 SCHPO
ID FAT1 SCHPO STANDARD; PRT; 1385 AA.
AC O13735; Q9UNJ0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churchman C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Murrell K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vansteenkiste E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
RN [2]
RP IDENTIFICATION, AND GENE NAME.
RX MEDLINE=20143585; PubMed=10679021;
RA Jin H., Amberg D.C.;
RT "The secretory pathway mediates localization of the cell polarity
RT regulator Alp3p/Bud6p.";
RL Mol. Biol. Cell 11:647-661(2000).
CC -!- FUNCTION: Involved in the organization and/or function of the

```

```

CC actin cytoskeleton.
CC -!- SIMILARITY: TO YEAST BUD6.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z97208; CAB10112.1; -.
DR EMBL; AL109770; CAB52420.1; -.
DR GeneDB SPombe; SPAC15A10.16; -.
DR GO; GO:0005884; C:actin filament; ISS.
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR InterPro; IPR005613; AIP3.
DR Pfam; PF03915; AIP3; 1.
KW Coiled coil; Cytoskeleton.
FT DOMAIN 1009 1096 COILED COIL (POTENTIAL).
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40CC4FF7537 CRC64;

Query Match 4.3%; Score 162; DB 1; Length 1385;
Best Local Similarity 19.2%; Pred. No. 1.2;
Matches 147; Conservative 125; Mismatches 293; Indels 202; Gaps 33;

QY 26 DLNQAPMWTSTTGL-----STPSSLENIIPSENQYFQSAIWS-----GFIVKXKS 74
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 326 DRKFSPLKTRTPSLTKSLDPTSLKSPSLRKSPS--SFVQKDVYSRNSLRISQANRS 383
QY 75 DEYFATSAADNHVTMVDQEVINKAS-----NSN 104
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 384 NVFPGATD---VTRSVSDHRLSSSTINDGEVAPLPQRSRTISSPNSPLSATVPSST 439
QY 105 KIRLEKRLYOIKIYQOREN-----PTEKGLDPKL-----YWTDSQNK 143
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 440 PILLGRGSRSTLSVKNKQFNADGDTLNSPNSIRETEEYAAAPKLEDIADDEVTDATSR 499
QY 144 EVI-----SSDNLQLPELKOKSS---NSRKKRSTSGAPVVPORDNDGI-----PDSL 187
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 500 ELLERQTKAESDSTSEIESLQKLSLPQVSTQEIQPSVSSVPEASNEIAEKPAVT 559
QY 188 EYEGYTVVDVKNKRTFLSPWT-SNIHEKKGLPKYKSSPEKWTASDPYDFPKVTGRDKNV 247
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 560 AIESIT-----ERKEEAPVISSEKIESGTSIGTSDTKGGLANFENDSLBELRLIQQN- 613
QY 248 SPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSRTHTSEVHGNA 307
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 614 NAEQDEPSYKPKYKYESEER-----SGSEDEFKSEKDTKGVYINDSDSTQVEESEDKSTP 669
QY 308 EVHASFFDIGGSVSAGFSNSNSTVAID-----HSLSLAGERTWAETMGLNTADTARLNAN 363
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 670 NTGA-----SAKLNDPSSITITVSDVYKPKASPVEITEPPSSALVSATSPTTNVP 720
QY 364 I--RYVNTGTAPYVNVLTPTSLVLGKQNTLATIKADENQLSOLAPNPNYPSKNLAPIAL 421
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 721 IVPEAVHLSTA--FSTAPVSTIV-----SNISPLPTVAP--P 753
QY 422 NAQKQDASSTPTIM--NTNQFLELEKT--KQLRLDTQVYGNIAFYFENGVRVYDTGSNW 477
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 754 NVSGSPSETPIKSEKPKYVVSQTEKALPKPLGVDTKEY--FLRYNNQTKKVKVSPSLN 811
QY 478 SEVLPOQETTARIIFNGK--DLNLVERRIAAVNPSPDLETTQPDWT-LKEALKIAPGFN 534
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 812 ANELGELFSNVYKISFGSDSYELNIED-----PDTKISYLLLEDLSLKSLVSPFMFK 864
QY 535 EPNGNLQ---YQKDIITEFDNFQDQTSQNIKNQLAEINATNIYT-VLDIKLNKAKNII 590
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 865 EQDANKKREDPFSGEVSAIGHSSAQNTLDDHVNTTHESPSSAFTLELELKA----- 917
QY 591 IRDKRFHYDRNNIAGADESVVKEAREVINSSTEGLLNIDKDIRKILSGYIV-EIEDT 649
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 918 -----IEQNIST-----NHTNDSAAKKSDESHSKLANNFVSPDSIDH 954

```

```

QY 650 EGLKEVINRDMNLNLSRLQDQKTFI-----DFKK-----YNDK 694
ID SMC3_YEAST STANDARD; PRT; 1230 AA.
Db 955 KFYQVQNMQ---LELASLKQISAFFTRPLKIKDFKKEINAFNEK 998

RESULT 11
SMC3_YEAST
AC P47037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Structural maintenance of chromosome 3 (DA-box protein SMC3).
GN SMC3 OR YUL074C OR J1049.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_taxID=4932;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=9303;
RX MEDLINE=97474309; PubMed=9335333;
RA Michaelis C., Ciosk R., Nasmyth K.;
RT "Cohesins: Chromosomal proteins that prevent premature separation of
RT sister chromatids.";
RL Cell 91:35-45(1997).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=96208490; PubMed=8641269;
RA Galibert F., Alexandraki D., Baur A., Boles E., Chalwatzis N.,
RA Chuat J.-C., Coster F., Cziepluch C., De Haan M., Domdey H.,
RA Durand P., Entian K.-D., Gattus M., Goffeau A., Grivell L.A.,
RA Hennemann A., Herbert C.J., Heumann K., Hilger F., Hollenberg C.P.,
RA Huang M.-E., Jacq C., Jauniaux J.-C., Katsoulou C., Kiehl J.,
RA Kleene K., Kordes E., Kotter P., Liebl S., Louis E.J., Manus V.,
RA Mewes H.-W., Miosga T., Obermaier B., Perea J., Pohl T.M.,
RA Portetelle D., Pujol A., Purnelle B., Ramezani Rad M., Rasmussen S.W.,
RA Rose M., Rossau R., Schaeff-Gerstenschlaeger I., Smits P.H.M.,
RA Scarcez T., Soriano N., Tovan D., Tzeremia M., Van Broekhoven A.,
RA Vandenbol M., Wedler D.H., Von Wettstein D., Wambutt R., Zagulski M.,
RA Zollner A., Karpingner-Hartl L.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT X.";
RL EMBO J. 15:2031-2049(1996).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1678;
RA Sor F.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
IDENTIFICATION IN A COHESIN COMPLEX WITH SMC1, IRR1 AND MCD1, AND
RP INTERACTION OF THE COHESIN COMPLEX WITH SCC2.
RX MEDLINE=99145468; PubMed=9990856;
RA Toth A., Ciosk R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;
RT "Yeast cohesin complex requires a conserved protein, Ecolp(Ctf7), to
RT establish cohesion between sister chromatids during DNA replication.";
RL Genes Dev. 13:320-333(1999).
RN [5]
IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3, MCD1 AND IRR1, AND
RP STRUCTURE.
RX MEDLINE=21980168; PubMed=11983169;
RA Haering C.H., Loewe J., Hochwagen A., Nasmyth K.;
RT "Molecular architecture of SMC proteins and the yeast cohesin
RT complex.";
RL Mol. Cell 9:773-788(2002).
CC -!- FUNCTION: Involved in chromosome cohesion during cell cycle and in
CC DNA repair. Central component of cohesin complex. The cohesin
CC complex is required for the cohesion of sister chromatids after
CC DNA replication. The cohesin complex apparently forms a large
CC proteinaceous ring within which sister chromatids can be trapped.
CC At anaphase, the complex is cleaved and dissociates from
CC chromatin, allowing sister chromatids to segregate.

```

```

CC -!- SUBUNIT: Cohesin complexes are composed of the SMC1 and SMC3
CC heterodimer attached via their hinge domain, MCD1/SCC1 which link
CC them, and IRR1/SCC3, which interacts with MCD1. The cohesin
CC complex also interacts with SCC2, which is required for its
CC association with chromosomes.
CC -!- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin.
CC Before prophase it is scattered along chromosome arms. At
CC anaphase, the MCD1 subunit of the cohesin complex is cleaved,
CC leading to the dissociation of the complex from chromosomes,
CC allowing chromosome separation.
CC -!- DOMAIN: The flexible hinge domain, which separates the large
CC intramolecular coiled coil regions, allows the heterotypic
CC interaction with the corresponding domain of SMC1, forming a V-
CC shaped heterodimer. The two heads of the heterodimer are then
CC connected by different ends of the cleavable MCD1 protein, forming
CC a ring structure (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. SMC3 subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y14278; CAA74655.1; -
CC EMBL; Z49349; CAA89366.1; -
CC EMBL; X88851; CAA61313.1; -
CC FIC; S56850; S56850.
CC GenOnline; 141688; -
CC SGD; S0003610; SMC3.
CC GO; GO:0007151; P:spoolation (sensu Saccharomycetes); IMP.
CC GO; GO:0007130; P:synaptonemal complex formation; IMP.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR003405; SMC C.
CC InterPro; IPR003395; SMC N.
CC Pfam; PF02483; SMC C; 1.
CC Pfam; PF02463; SMC N; 1.
CC Mitosis; Cell cycle; Chromosome partition; ATP-binding; Coiled coil;
CC Nuclear protein.
CC NP_BIND 32 39 ATP (POTENTIAL).
CC DOMAIN 172 482 COILED COIL (POTENTIAL).
CC DOMAIN 483 684 FLEXIBLE HINGE.
CC DOMAIN 685 1041 COILED COIL (POTENTIAL).
CC DOMAIN 1126 1161 ALA/ASP-RICH (DA-BOX).
CC SEQUENCE 1230 AA; 141336 MW; B152D88F780341F CRC64;
CC
Query Match 4.3%; Score 161.5; DB 1; Length 1230;
Best Local Similarity 19.7%; Pred. No. 1;
Matches 181; Conservative 136; Mismatches 356; Indels 247; Gaps 38;
QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVT--SSTTGDLSIPSSSELENIPSEHQ 58
Db 209 EMEQERKELEKYNELERNKIKYQFTLYDRELNEVINQMERLDGYN-----NTVYSSSQ 262
QY 59 YFQ-----SAIWSGPIKVKSDVTFATSDNHTVMVDDDEVINK---- 99
Db 263 YIQELDKREDMIDQVSKLSSI-EASLTKNATDQAKLRESEISQKLTNNVNIKDVQ 321
QY 100 ---ASNSNKIRLEKGRLOYK-IQYORENPTKGLDFKLYWTDSONKKEVISDNLQLE 155
Db 322 QQIESNEBORNLDATLKEIKSIIEQKOKLSKILPRYQELT-----KEAMY--KLQLAS 375
QY 156 LKQKSSNRKRSSTAGPTVPDRNDGIPDSLEVEGYTVDVKNKRTFLSPWTSNTHE-KK 214
Db 376 LQOKQDLTLKKEGEARFKSKDERDTWIHSEI-----ELKSS-----IQNLNELES 422
QY 215 GLTKYKSSPEKSTASDPYSDFEKTGTGRDKNVSPEARHPLVAAPIVHVDMENILSKN 274
Db 423 QLQMDRTSLRKQVSAID-----EEIEELDSINGDPTKGQL-----EDFSELHLKQK 471
QY 275 EDQSTQNTDSETRTTSKNTSTSRITSEVHGNAEVSASFDDIGGSVAGSAGSNSSTVAI 334

```

Db 472 LSESLDTRKELWRKEQLQTVLEILLSVQNR-----NNVTMSRSLANGIINVKEL 525  
Qy 335 DHSLSLAGE-----RTWAETMG-----LNTADTARLNANIRVYNTG 370  
Db 526 TEKIKISVESVFGILGELIKNDKYKTCAEVIGNSLFHIVVDTEETATLIMNELYRMKG 585  
Qy 371 TAPIYVLPVTTSLVLGKNGQTIATIKADENQSQILAPNNYPSKXLAPIALNAQKDSST 430  
Db 586 GRVTF--IPLNRLSLSDVKFSPSTTTQIQTPLIKIKYEPFEKA-----VKHVPKGT 638  
Qy 431 PITVYNYNPFLEKTKQLR--LDDQV-YGNAT--YNFENGVRVDTGNSKSEVLPOI 484  
Db 639 IVVKDLGGGLAKXKHUNATLDGRADKRGVLTGGYLDOKKTRLESKLNLSRSQH 698  
Qy 485 QETTARIIFNGKDLNLVERRIAANPS-----DPLETKPD-MTLK 524  
Db 699 KKILEELDFVRNELNIDTKIDQVNGNIRKVSNDRESVLTIWVVRTSLNKKNEKLE 758  
Qy 525 EALK-IAGFNEPNGNLQYQKDIETDFN-----FDQOTSQNIKNOLA-----E 568  
Db 759 ESLNAILIKLEKLNTRTFAOEKNTFENDLLOQFDSLSKEERLESUTKEISAHNK 818  
Qy 569 LNATN-----IYVLDKIKLAKMMLIRKRFHYDRNNIAG-----A 607  
Db 819 LNITSDALEGITTTID--SLNAELSKLIPQENDLESKMSEVGDATFGLQDELKEIQLE 876  
Qy 608 DESVKEHAREVINGST-----EGLL-----NIDKDIRKI 638  
Db 877 KESVEKOHENAVLELGTVQREIESIABETNNKLEKANNOORLLKKLDFNFKSVEKT 936  
Qy 639 LSGYIVETEDTGLKE-----VINDRYDM-----LN-----ISSLRQDGK 673  
Db 937 MIKKTLVTRRELQORREICLLPEDALVDFSDITSDQLLORLNDMNTWISGLKNVK 996  
Qy 674 -TFIDFKKNYK-----LPLYSNPNYKNVAVYKNTIINPSENG 714  
Db 997 RAFENFKFENRRKOLAEASELDESQSDIOLIVKLKQKYNV-----DSTFKQVSENF 1052  
Qy 715 DT-----STNGIKKILIFSK 729  
Db 1053 EAVFERLVRPGTAKLIHRK 1072

RESULT 12  
P115 MYCHR  
ID P115 MYCHR STANDARD; PRT; 979 AA.  
AC P41506;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE P115 protein.  
OS Mycoplasma hyorhinitis.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2100;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=91138990; PubMed=1825306;  
RA Notaricola S.M., McIntosh M.A., Wise K.S.;  
RT "A Mycoplasma hyorhinitis protein with sequence similarities to  
RT nucleotide-binding enzymes.";  
RL Gene 97-77-85 (1991).  
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -I- DOMAIN: Consists of two putative central coiled-coil regions  
CC flanked by putative globular regions at the N- and C-termini.  
CC -I- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA  
CC P115.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M34956; AAA25423.1; .  
DR PIR; JQ0894; JQ0894.  
DR InterPro; IPR003439; ABC transporter.  
DR InterPro; IPR005289; GTP-binding\_dom.  
DR InterPro; IPR003405; SMC C.  
DR InterPro; IPR003395; SMC N.  
DR Pfam; PF02483; SMC C; 1.  
DR Pfam; PF02463; SMC N; 1.  
DR TIGRfams; TIGR00650; MG442; 2.  
KW ATP-binding; Coiled coil. ATP (POTENTIAL).  
FT NP\_BIND 32 39 COILED COIL (POTENTIAL).  
FT DOMAIN 169 244 COILED COIL (POTENTIAL).  
FT DOMAIN 231 400 COILED COIL (POTENTIAL).  
FT DOMAIN 569 821 COILED COIL (POTENTIAL).  
FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).  
SQ SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;

Query Match 4.3%; Score 160.5; DB 1; Length 979;  
Best Local Similarity 19.0%; Pred. No. 0.86;  
Matches 176; Conservative 132; Mismatches 342; Indels 275; Gaps 40;

Qy 13 ESSSQGLGYPFSDLNFOAPMVV-----TSSTTGDLSIPSELENIPSE 56  
Db 51 EQSAQLRLGNMDDVIFAGSKTVKQKAMVYKTFKNEDAIEETKQIFITSLLRGQGT 110  
Qy 57 NQYFQS-----AIVSGFIKVKSDRYTATSADNHTVMVDDOE--VINKASN 102  
Db 111 NEFYNDQPVRYNDIKNLAVESG---ISKSLAISQGTISEIABATPEQRKAVIEEAG 167  
Qy 103 SNKIRLEKRLYQIKIYQORENPTEKGLDFKLYWTDSQNKKEVSSDNLQLPELKQKSN 162  
Db 168 TSKYKLDKEEAQKLR-----TNDALD-KLQGAIKELERQVNSLDK-QASKAKIYLEK 219  
Qy 163 SRKRRTSAGTPVPRD-----NDGIPDSLEVEGYTVDVK-NKRTFLSPWISNHEKKG 215  
Db 220 SKALESVEVGLIYNDLNFNEKLNLTSLLEVEQORNDLELNIQTYESSISQTVH---- 275  
Qy 216 LTKYKSSPEKMWSTASDPYSDPEKVTGRID--KNVSP-----EAPPLVAAPVIVHV 264  
Db 276 ---FKTEV-----SSIQEITSKLDNLKALSEINLQEARIEERKLIISGEIV-V 322  
Qy 265 DMENII--LSKN-EDOSTQNTDSETTISKNTSTSTHT----- 300  
Db 323 DQKTKIEEIKQVESLKIQINASKQREIELDQOLTRLNAKANSKLQENDINKEIGVLLE 382  
Qy 301 --SEVHGNAEVHASFFDIGSVSAGFSNSSTVAIDHSLSLAGERTWAB---TWGLNTA 355  
Db 383 KKSAAAAANILKQOFENKSFSGIKTKIDNSFLPDGYIGLASELKFVSESPSLAIEIV 442  
Qy 356 DTARLN-----ANIRVYNTGTAPIYVLPVTTSL-----VLG 386  
Db 443 LGAALNQIVMKTSEDVLQALDFLKNLSGKATFIPLTISIKEREVRDHLVLKQKGFGL 502  
Qy 387 KNOTLATIKADENQ-----LSQILAPNNYPSKXLAPI-----ALNAQKDSSTPT 433  
Db 503 VAKELIEFTDQFNKLFGLIGNILVVDNVDNANRIAKILDHKYTVLSLGGDIFRPGGTIT 562  
Qy 434 -----MNYNQFLELEKTKQLRLDQ-----VYGNATYFNFENGVRVDT 473  
Db 563 GGSKLRTSLNLYDIKIK-EHTNLTAFABEDQIHLKIKQOTTIYNEIETVNSTIQQVKIEA 621  
Qy 474 GS-----NWSVLVQIQE--TTARIIINGKDLNLVERRIAAV-----N 509  
Db 622 NSINSKNLINLNEBLNKLNASEIFKEQOBDQBSLNSLDFSEKLN-IEQKISTLTTELNS 680  
Qy 510 PSDDL-----ETTKPDMTLK-----EALKIAGFNEPNGNLQYQGD 546  
Db 681 KQRLTNLISEQKGETKQELDAKLKUNTOHSDSITSEQNRAKFLVEQNKELSEHYKL 740  
Qy 547 ITE-----PDFNFDQOTS--QNIKNQALNATNIYVLDKIKLNKAKNMLIRKRFH 597





```
QY 55 SENQYFQSAIWGFIK--VKKSDYTFATSDNHVTM----- 89
Db 329 TTNYSASSQKWDANLXAYLYLKTGYD-KTTGTNYAFMKFNLLKPTQNMVTKATLKYVA 387
QY 90 -----WYDDQEVINK-----ASNSKIRLEKGR--LYQIKIQYOR 122
Db 388 HSYGTYKATGLMLD---TVNSYNDNAKVNTWTKPASKNIGADVHKQWASDYDTAAVKS 444
QY 123 ENPTEKGLDFKLYWTDSONK---KEVSSDNLQLPELKQKSNRKKRSTAGTVP--- 176
Db 445 WNSGGANYGFKLH--TNGNKEKYWKLLIS-----SANSANKPIYEVITTIKGN 491
QY 177 -----DRDNDGIPD---SLE-VEGYTVDVKNKRTFLSPFWISNHEKGLTKYKSSP 223
Db 492 TPTIKAYHNGDSTGYFDISWKKVEGAGYKWIYNGKEYQALSAGNV-----TSWSTKG 545
QY 224 EK-WSTASDPYDFEKVTRID-KNVSPARPLVAAYPIVHVDMENTILSKNE----- 275
Db 546 KKLWTSABIAIKRYKL--HLGDKGAEALADP---SPVYKNGSGSYATSKNTWIGVSA 599
QY 276 --DQSTQNTDSETRTISKNTSTSRTHSEVHGNAEVAHSPFDIGGSVAG-----FS 325
Db 600 IFDQEGGASAPAKPVPVWVKAQAPSAKGYNGNA--TGYPDLWKAVSGATGYKVQVFN 658
QY 326 NNSNSTVAIDH--SLSLACERTWETMGLNTADTA-----RLNANIRVYNTGTA 372
Db 659 GKGFTLDLGNQTSWTKKLIWPTSAEIKAGYALHLKDGSGAELPINTPGTYKNAGGD 718
QY 373 PIYNVLFTSLVGLKQNTLATIKADENQLSQILAPNN-----YVPSK-----NLA--- 417
Db 719 GAKRNSYFKIIAYNKDGEALASPAATPALPIARPKNTGYLYNTKSSQGYVNLWEK 778
QY 418 -----PIALNAQKASSTPI-----TMNYNQFLELETK-----QJRL 450
Db 779 VQNAKGYKVIYNGKEYQFSDGADHWTTQNKNIWPTSEIKAGSYKXLTGDKGGELAL 838
QY 451 DTDQVYGN-----IATYN--FENGVRVDVTSNWSV-L 481
Db 839 DPSPVYNNANGYKKNYSFTLVAYDANGETIPTAPENPTFHEGAELFTEEYWSIIDI 898
QY 482 P--QIQTETARIIFNGKDLNL-----VERIAAVNPSDPL-----ETT----- 517
Db 899 PSGLNGATGNVIVNEEDLISIDRGFGGLSTYNSLSDHFLFGQGWYDAETSVISTD 958
QY 518 -----KPDWTLKEALKIAG-FNBPNGNLQYQGDITEFDNFDDQTSQNIKNQLAEL- 569
Db 959 GGAWYDEDATTHRFPKADGTGYPTGYVLELTETADQFILKTQDQINAYFNKKGKQLQ 1018
QY 570 -----NATNIVTLDKIKLNAKMILLIRDKRFHYDRN-----NIAVGADES 611
Db 1019 KVVGDHNNATVYNDKNQLTAITDASGRKLTFTYDENGHVTSITGPKXKKVYTSYENDL 1078
QY 612 VKE-----RKLSCY-----IVEIEDTEGLKXVINDRYDMLNI 665
Db 1079 LKKVTDGTGVTSDYDSEGLRVKQVYNSANSTEAKPVFTEYQYSGHRLKAINAKKETVY 1138
QY 630 NIDKDI-----RKILSY-----AHR--EVINSSTEGLLL 629
Db 1139 SYDADKXLLMTQPNRKYQVGYNEAGNPQIOWIDDAEGLKITNTKYEGNV 1190

RESULT 15
CAGA_HELPJ
ID CAGA_HELPJ STANDARD; PRT; 1167 AA.
AC Q3ZLTI;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytotoxicity associated immunodominant antigen (120 kDa protein)
DE (CAGA pathogenicity island protein 26).
GN CAGA OR CAI OR CAG26 OR JHP0495.
CS Helicobacter pylori 399 (Campylobacter pylori J99).
OC Bacteria, Proteobacteria; Epsilonproteobacteria; Campylobacterales;
```

Helicobacteraceae; Helicobacter.

NCBI\_TaxID=85963;

[1]

SEQUENCE FROM N.A.

MEDLINE=99120557; PubMed=9923682;

Alm R.A., Ling L., S.L., Moir B.T., King B.L., Brown E.D., Doig P.C.,

Smith D.R., Noonan B., Gullid B.C., deJonge B.L., Carmel G.,

Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,

Gibson R., Metzberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,

Trust T.J.;

RT "Genomic sequence comparison of two unrelated isolates of the human

gastric pathogen Helicobacter pylori.";

Nature 397:175-180(1999).

CC FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,

OR FUNCTION OF THE CYTOTOXIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB001483; AAD06073.1; -.

PIR; B71924; B71924.

InterPro; IPR005169; Caga.

InterPro; IPR004355; IVSec\_caga.

Pfam; PF03507; Caga; 1.

PRINTS; PR01553; TYPE4SSCAGA.

Antigen, Complete proteome.

DOMAIN 246 249 POLY-THR.

DOMAIN 882 889 POLY-ASN.

SEQUENCE 1167 AA; 129729 MW; FD5E6B981CEBD0F2 CRC64;

Query Match 4.2%; Score 158.5; DB 1; Length 1167;

Best Local Similarity 20.5%; Pred. No. 1.4;

Matches 167; Conservative 138; Mismatches 297; Indels 213; Gaps 44;

QY 3 KOENPLLESSESSOGLL---GYFSDL-NP-----QAPKVTSTTGLSLPSSELE- 51

Db 480 KADKALDREAKTTQGLSKHGVNPFVSNPKYTNASKSPDKGVGATNG-----VSHLEA 534

QY 52 -----NIPSENYFQSAIWGFIKVKSDYTFATSDNHVTWVDD-----QEVIN 98

Db 535 GFSKVAVENLNLNLAITSV---VRQDLEDKLIAGLSPQEAANKLVKDFLSSNKELVG 590

QY 99 KASNNKTRLE---KGRLYQIKIQYQRENTEKGLDFKLYWTDSONKKEVISSDNLQLOPE 155

Db 591 KALNFNKVAEAKTNGYDEVK---QACKOLEKSL-----KKRERLEKDVAK--N 635

QY 156 LKQKSSNRKSRKSTISAGTVPDR-----DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNI 210

Db 636 LESKSGNKNKMEAKSQANSQKDEIFALINKEANRDARAI-AYAQNLIKIKRELSDKLENI 694

QY 211 HEKKGTLTKYKSPKSWSTASDPYSDPEK-----VTGRI-DKNVSPARHPLVAAYPIV 262

Db 695 N--KDLKDFSKSFDEFKNGKN--KDFSKASETLKALGSKVDLGINPEWIS----- 741

QY 263 HVDMENIILSKNEPQSTQNTDSETRTISKNTSTSRTHSEVHGNAEVAHSPFDIGGSVSA 322

Db 742 --KVENLNAALNEFPNGKNKDFSKVTOAK--SDLENSIKDVIINQKITDKVDNLNQAVSV 797

QY 323 GFSNNSNSTVALDHSLSLAGERTWAEWTMGLNTADTARLNANIRVYNTG-TAPIYVNLPTT 381

Db 798 AKATGDFSGV-----EQALADLKNFSKEQLAQQAQKQNDFTNGKNSALYQ----- 842

QY 382 SLVLGKQNTLATIKADENQLSQILAPANNYPSKNLAPI--ALANQ-----KDS 428

Db 843 SVKNGVNGTLV-----GNLSKAEATT---LSKNFSDIKKELNAKLGNFNNNNNGLNS 894

QY 429 STPIWNTNQFLELETKQLBLDQVYGNIAATYTNFENGVRVDVTSNWSVLPQJQETT 488

Db 895 TEPI---YTQVAKVKAKIDRL--DQIASGLG-----DVQOASFLLRHDKV- 937  
Qy 489 ARIIFNGKDLNLVERRIAAVNPSPLETTKPD-----TLXEALKIARGFNEPENGNLQYOG 544  
Db 938 -----DDL-----KVGLSANHEPIYATIDDLGGPPLXRHDKV----- 971  
Qy 545 KDI TEFDNFQOTSQNIKN---QJAEINAT---NIYTVLDKIKLNAKMIILIRDKRFHY 598  
Db 972 DDLSKVGLSREQKLTQKIDNLNQAVSEAKASHFDNLDQMIDKLDKSTKKVW----- 1023  
Qy 599 DRNNIAGADESVVKEAHREVINSSTEG---LILNI-DKDIRKILSGYIVEIEDTEGLKE 654  
Db 1024 ---NLYVESAKVPTLSAKLDNATNSHTFEINSVXNGTINEKATCMLTQ-KNSEWLK- 1078  
Qy 655 VINDRYDMLNISSLRQDGKTFIDF-----KKYND--KLPLYISNP-----NYKVNV 698  
Db 1079 LVNDKIVAHNVGSAPLSAYDKIGFNGKMKDYSDSPKFSFTRLSNAVKDIKSGFVQFLTNI 1138  
Qy 699 YAVTKENTIIINPSENGDSTNGIKILIFS KGYE 733  
Db 1139 FSMGSYSLMKASVEHGVKNTN-----TKGGFQ 1165

Search completed: May 3, 2004, 19:37:04  
Job time : 13.2539 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds  
(without alignments)  
6697.424 Million cell updates/sec

Title: US-09-848-909A-13  
Perfect score: 3772  
Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKKILFSSKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rviro.\*  
16: sp\_bacterioph.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3681	97.6	741	2 Q937W2	Q937W2 bacillus an
2	3681	97.6	743	2 Q937W3	Q937W3 bacillus an
3	874.5	23.2	876	2 O32739	O32739 clostridium
4	874.5	23.2	879	2 O06498	O06498 clostridium
5	872.5	23.1	876	2 Q9KH41	Q9KH41 clostridium
6	868	23.0	875	2 Q46221	Q46221 clostridium
7	810	21.5	721	2 O86171	O86171 clostridium
8	720	19.1	775	2 O844J8	O844J8 bacillus th
9	296.5	7.9	225	2 Q8YK2	Q8YK2 bacillus an
10	222.5	5.9	1881	16 Q8RGK2	Q8RGK2 fusobacteri
11	215.5	5.7	3468	5 Q8II04	Q8II04 plasmodium
12	202	5.4	4688	16 Q9PQ08	Q9PQ08 ureaplasma
13	200.5	5.3	2849	5 Q8IHY4	Q8IHY4 plasmodium
14	198.5	5.3	1904	5 Q8IKS2	Q8IKS2 plasmodium
15	197.5	5.2	2940	5 Q8IHP9	Q8IHP9 plasmodium
16	196	5.2	1404	2 Q8RJN9	Q8RJN9 mycoplasma

17	193.5	5.1	675	16 Q897H6	Q897H6 clostridium
18	192.5	5.1	2771	5 Q26216	Q26216 plasmodium
19	191.5	5.1	2444	16 Q8IGX1	Q8IGX1 bacillus ce
20	190	5.0	9904	16 Q8NWQ6	Q8NWQ6 staphylococ
21	189.5	5.0	2529	16 Q25579	Q25579 helicobacte
22	189	5.0	3248	5 Q8ICP9	Q8ICP9 plasmodium
23	188	5.0	2792	5 Q8I4R2	Q8I4R2 plasmodium
24	187.5	5.0	1301	5 Q8WSK5	Q8WSK5 plasmodium
25	187.5	5.0	1681	5 Q8I548	Q8I548 plasmodium
26	186.5	4.9	1125	16 Q98PM9	Q98PM9 mycoplasma
27	186.5	4.9	2235	5 Q8IB09	Q8IB09 plasmodium
28	186	4.9	1859	5 Q8IC27	Q8IC27 plasmodium
29	186	4.9	3504	5 Q8IL45	Q8IL45 plasmodium
30	186	4.9	4261	5 Q8IFP4	Q8IFP4 plasmodium
31	183	4.9	3203	5 Q7YWE7	Q7YWE7 plasmodium
32	183	4.9	3203	5 Q7YWE6	Q7YWE6 plasmodium
33	183	4.9	6713	16 Q99U54	Q99U54 staphylococ
34	183	4.9	6713	16 Q931R6	Q931R6 staphylococ
35	183	4.9	10061	5 Q8I3Z1	Q8I3Z1 plasmodium
36	182.5	4.8	1113	5 Q8ISQ0	Q8ISQ0 plasmodium
37	182.5	4.8	1301	5 Q8IHQ2	Q8IHQ2 plasmodium
38	182.5	4.8	2049	5 Q8IDA9	Q8IDA9 plasmodium
39	182	4.8	2747	5 Q9BJX9	Q9BJX9 plasmodium
40	181.5	4.8	1263	5 Q8IK49	Q8IK49 plasmodium
41	181.5	4.8	1387	5 Q9GZ76	Q9GZ76 plasmodium
42	181.5	4.8	3080	5 Q7YWF0	Q7YWF0 plasmodium
43	181	4.8	1151	5 Q8ILU0	Q8ILU0 plasmodium
44	181	4.8	1302	2 Q49547	Q49547 mycoplasma
45	181	4.8	2269	5 Q26223	Q26223 plasmodium

## ALIGNMENTS

RESULT 1

Q937W2 PRELIMINARY; PRT; 741 AA.  
AC Q937W2  
DT 01-DEC-2001 (TREMREL. 19, Created)  
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)  
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)  
DE Pag protein (Fragment).  
GN PAG.  
OS Bacillus anthracis.  
OG Plasmid pXOI.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ferrara;  
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,  
RA Pasanella A., Francia M., Ciuchini F.  
RT "Sequence analysis of the genes encoding for the major virulence  
RT factors of bacillus anthracis vaccine strain 'Carbosap'."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ413937; CAC93935.1; -  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
DR GO; GO:0015070; P:toxin activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax\_toxinB.  
DR Pfam; PF03495; Anthrax\_toxB; I.  
DR PRINTS; PRO1391; BINARYTOXINE.  
KW Plasmid.  
FT NON\_TER 1  
FT NON\_TER 741  
SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 97.6%; Score 3681; DB 2; Length 741;  
Best Local Similarity 99.3%; Pred. No. 5.3e-167;  
Matches 717; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 EVKQENLLNESSSQGLLYFYFDLNFQAPWVTSSTTGDLSIPSELENIPSENQYF 60



```
Db 20 EVKQENRLNSESQGLGYFSDLPAPMVVTSSTTGDLSIPSSLENIIPSENQYF 79
QY 61 QSAIWSGFTKVKKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 80 QSAIWSGFTKVKKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 139
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKRSTSGPTVDPDRN 180
Db 140 QRENTEKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKRSTSGPTVDPDRN 199
QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEYV 240
Db 200 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEYV 259
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 260 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 319
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
Db 320 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 379
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 380 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 439
QY 421 LNAQKDSASTPIIMVYNOQFLEKTKOLRLDQVYGNATYVNFENGRVRVDTGSNWSEV 480
Db 440 LNAQKDSASTPIIMVYNOQFLEKTKOLRLDQVYGNATYVNFENGRVRVDTGSNWSEV 499
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Db 500 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 559
QY 541 QYQKDIETEPDFNDQOSTQNIKNQLAELNATNIVTVLDKIKNAKNNILLIRDKRHFYDR 600
Db 560 QYQKDIETEPDFNDQOSTQNIKNQLAELNATNIVTVLDKIKNAKNNILLIRDKRHFYDR 619
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Db 620 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 679
QY 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINSENGDTSTNG 720
Db 680 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINSENGDTSTNG 739
QY 721 IK 722
Db 740 IK 741

RESULT 2
Q937M3 PRELIMINARY; PRT; 743 AA.
AC Q937M3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pag protein (Fragment).
GN PAG.
OS Bacillus anthracis.
OC Plasmid pXOI.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=Carbosap;
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Fasanella A., Francia M., Ciuchini F.;
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbosap'.";
RL Submitted (SEF-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL: A7413936; CAC93934.1; -.
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0046821; C:extrachromosomal DNA; IEA.
DR GO: GO:0015070; F:toxin activity; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR GO: GO:0003896; P:anthrax toxinB.
DR Pfam: PF03495; Binary toxB; I.
DR PRINTS: PR01391; BINARYTOXINB.
DR Plasmid.
KW NON_TER 1 1
FT NON_TER 743 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5D5A6 CRC64;

Query Match 97.6%; Score 3681; DB 2; Length 743;
Best Local Similarity 99.3%; Pred. No. 5.3e-167;
Matches 717; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVQENRLNSESQGLGYFSDLPAPMVVTSSTTGDLSIPSSLENIIPSENQYF 60
Db 22 EVQENRLNSESQGLGYFSDLPAPMVVTSSTTGDLSIPSSLENIIPSENQYF 81
QY 61 QSAIWSGFTKVKKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 82 QSAIWSGFTKVKKSDSEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 141
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKRSTSGPTVDPDRN 190
Db 142 QRENTEKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNSRKRSTSGPTVDPDRN 201
QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEYV 240
Db 202 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEYV 261
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 262 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 321
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360
Db 322 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 381
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 382 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 441
QY 421 LNAQKDSASTPIIMVYNOQFLEKTKOLRLDQVYGNATYVNFENGRVRVDTGSNWSEV 480
Db 442 LNAQKDSASTPIIMVYNOQFLEKTKOLRLDQVYGNATYVNFENGRVRVDTGSNWSEV 501
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Db 502 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 561
QY 541 QYQKDIETEPDFNDQOSTQNIKNQLAELNATNIVTVLDKIKNAKNNILLIRDKRHFYDR 600
Db 562 QYQKDIETEPDFNDQOSTQNIKNQLAELNATNIVTVLDKIKNAKNNILLIRDKRHFYDR 621
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Db 622 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 681
QY 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINSENGDTSTNG 720
Db 682 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIIINSENGDTSTNG 741
QY 721 IK 722
Db 742 IK 743

RESULT 3
O32739
ID O32739 PRELIMINARY; PRT; 876 AA.
```





```

RN  SEQUENCE FROM N.A.
RP  STRAIN=NCIB 10748;
RX  MEDLINE=94041637; PubMed=8225592;
RA  Perelle S., Gilbert M., Boquet P., Popoff M.R.;
RT  "Characterization of Clostridium perfringens iota toxin genes and
RL  expression in Escherichia coli.";
RN  Infect. Immun. 61:5147-5156(1993).
RN  (2)
RP  SEQUENCE FROM N.A.
RA  STRAIN=NCIB 10748;
RC  Popoff M.R.;
RL  Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR  EMBL; X73562; CAA51960.1; --
DR  F01; 140862; I40862.
DR  HSSP; P13423; IACC.
DR  GO; GO:0005576; C:extracellular; IEA.
DR  GO; GO:0015070; F:toxin activity; IEA.
DR  GO; GO:0009405; P:pathogenesis; IEA.
DR  InterPro; IPR003896; Anthrax toxinB.
DR  Pfam; PF03495; Binary toxinB.
DR  PRINTS; PR01391; BINARYTOXINB.
KW  Signal.
FT  SIGNAL 34 38 POTENTIAL.
FT  CHAIN 212 875 IOTA TOXIN COMPONENT IB.
SQ  SEQUENCE 875 AA; 98468 MW; C9AF092CD3818921 CRC64;

Query Match 23.0%; Score 868; DB 2; Length 875;
Best Local Similarity 31.1%; Pred. No. 2,7e-33;
Matches 252; Conservative 130; Mismatches 279; Indels 150; Gaps 31;

QY 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTGDLSPSELENIPSE-NQY 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 35 DTNQKEITNENTLSSGLMGYFADHFHFKOLELAPIKNGDLKFEKKVLDLFDN 94
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 60 FQSAIWGGTKVKKSDEYFATSDNHNVTWVDQDEVINKASNNKIRLEKGLYQIK 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 95 IKSIWTGRIIPSEDEYILSTR-NDVLQINAKGDIK---TLKVNKKQVANNRIE 150
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 YQREN-----PTEKGLDFKLYWTDSONKKEVISSDNLOLPELKQSSNRKRSTS 170
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 IQDKNLGSDNLGVP-----KLYW-ELGNKNTVPEENLFFRDYKIDEND----- 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 AGTVP-----DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIH 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 196 --PFIPNNFDFVRFSSAAWEDDLTDNDNPDAYEKNGYTI-----KDSIAVKVND 249
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 EKKGKTKYKSSPEKSTADPYSDFKVTGRIDKNVSPPEARPLVAAPIVHVDMENIL 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 250 E-QGYKKYSSYLENTAGDPTDYQKASGSDKALKLEARDPLVAAPVGVGHNLI 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 272 SKNEDOSTQNTDETTISKTSTSTSTSEVHGNNAEVHASFDDIGGSVSAFSSNSST 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 309 STNEHASS---DQKTVSRATTSKTDANTV-----GVSIAGYQNGFTGN 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 332 VAIDHS-----LSLAGERTWAETMGLNTADTARLNANIRVNTGTAPIYVLP 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 ITTSYHTTNTSTAVQDSNGESNWTGLSKNGESAYINANVYNTGTAPYKVTPTNL 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 384 VLKGNQTLATIKADENQLSQILAPNNYFSKNLAPIALNAQKADASTPTMYNNGFLE 443
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 412 VL-DGETLATIKADNQIGNLSPNETYPKGLSLPLALNTMQFNARLIPINDQLKKLD 470
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 444 KTKQLRLDQVYGNATYVNFENGRVVDVTSNWSSEVLPOIQTETARIIFNGKDLN 503
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 471 SGQIKLEITQVSGNGYTKN-SGQI-ITEGNSWNSYISQIDSVSASIIID-TGSQ 527
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 504 RIAAVNPSPLETTKPDMTLKEALKIAGFNPNPQNGNQY-OQKDLITE--FDNF 560
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 528 RVAAKEQGNPEDKT-FEITIGEAIKKAFASTK-NGELLYFNGIPIDESCVELIFD 585
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 561 NIKNQIAELNATNIYVLQIKLNAQONILIRDKRF--HYDR-NNIATVCADESUVK 617
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 586 IIKEQLKYLDDKKIYV----KLERGNLILKVPSYFTNPFDEYNNFP--ASW 639
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 618 EVINSSTEGT-----LLNDIDKDIRKILSGY-----IVEIEDTEGLKEVIN 658
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 640 DGLGVANKLSGETKIIPMSKLPKRYVFGSKDPSTNSITVNIKSKQKTDYLPV 699
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 659 RYDMLNLS-----SLRQDKGTFFDFKKYND--KLPLVISPNPKV----- 696
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 700 EKDYTFSEYFETGKDDSDIEITLTSSGVIFDNLISITELNSTPILKEPIKVP 759
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 697 -----NVYAVTKENTIINSENGDTSTNGI 721
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 760 ILDAHNKYADIKLDT-----NTGNTYIDGI 785
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
O86171 AC O86171 PRELIMINARY; PRT; 721 AA.
ID AC O86171;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE C2 toxin (Component-II).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=96323874; PubMed=9659689;
RA Kimura K., Kubota T., Ohishi I., Isogai H., Isogai B., Fujii N.;
RT "The gene for component-II of botulinum C2 toxin.";
RL Vet. Microbiol. 62:27-34(1998).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=96184657; PubMed=8645309;
RA Fujii N., Kubota T., Shirakawa S., Kimura K., Ohishi I., Moriishi K.,
RA Isogai E., Isogai H.;
RT "Characterization of component-I gene of botulinum C2 toxin and PCR
RT detection of its gene in clostridial species.";
RL Biochem. Biophys. Res. Commun. 220:353-359(1996).
DR EMBL; D88982; BAA32537.1; --
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 721 AA; 80515 MW; 44C8153AC749DSF2 CRC64;

Query Match 21.5%; Score 810; DB 2; Length 721;
Best Local Similarity 35.3%; Pred. No. 1,2e-30;
Matches 223; Conservative 117; Mismatches 228; Indels 64; Gaps 25;

QY 3 KOENRLNES--SSQGLGYFSDLNFOAPMVVTSSTGDLSPSELENIPSENQYF 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 5 KFENSVKNSKNRYFTINGLMGYFEN-DFNLIISITPLDGNLTFKEDINSILG-NKII 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 QSAIWSGFIKVKKSDEYFATSDN-HVTWVDQDEVINKASN-SNKIRLEKGLYQIKI 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 KSAEWIGLIKPSITGEVILSTNSPNCVEL--NGEIFNLSTNTNTVNLIQGVYDIRI 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 -QYORENPTKGLD-FKLYWTDSONKKEVISSDNLOLPELKQSSNSRKRSTSAGTV- 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 120 EQLMSENQLLNKNGIKLYWETSDIIEIIPSEVL-----LKNYSNTNEKSFIPN 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 -----PDRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEK 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 176 SNAKLKANARNDTRDGIPEWEINGYTVNQKAVAWDDKFAAN-----GYK 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

226 WSTASDPYDPEKVTGRIDKKNVSPPEARHPLVAAPYPIVHVDVMEIILSKNEEDOSTONTDSE 285  
231 PCTANDPYDPEKVSQIDPSVSWARPMISAYPIGVQWERLVVSKSE-----TITGDS 286  
286 TRTISKNTSTSTHTSEVHNAEVAHSPFDIGG-----SVSAGFSNSNSTVAIDHSL 338  
287 TKSMKSTSHSNTINTV--GAEVSGSLQAGGIPPFVFSMSASANSYHTWQNTSTVDDT- 343  
339 SLAGERTWAEWTGLTADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADE 398  
344 --TGE-SFSGSLINIGSAYINPIRYNTGTAPVYNTPTTIVIDK-QSVATIKQE 399  
399 NOLSOILAPNNYPSKNLAPIALNAOKDASSTPIITMYNQFLEKTKQLRLDQVYGN 458  
400 SLIGDYLNPFGGTYPIGEPMALNTMDOFSRLIPINYNQLKSIDNGGTVMLSQFTGN 459  
459 IATYFENGVRVVDGNSWSEVLPOIQTETARII--FNGKDLNVERIAAVNPSDPLET 516  
460 FAKYM-SGNLIVTD--GNWGPVLGKTKTASLTLSFSGQTTQVA--VVAFNPSDEBK 514  
517 TKPDMTLKEALKIARFNEPNNGNLOVQKDIIT--EFDNFDOQTSQNIKNQLAELNATN 573  
515 T-PKLTLEQALVAFALKKNGKVFYHGLEISKNEKIQVFLDSNTNDFENQLKNTADKD 573  
574 IYVLDKIKLNAKMMILLIRKPHYDRNNIAV 605  
574 IMHCI--IKRN--MNLVKVITFKENISSINI 601

RESULT 8  
Q844J8  
ID Q844J8 PRELIMINARY; PRT; 775 AA.  
AC Q844J8  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Viplac.  
OS Bacillus thuringiensis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1428;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shi Y., Chen J., Pang Y.;  
RT "Cloning of vip1A(C) and vip2A(C) from Bacillus thuringiensis HD201.";  
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY245547; AA086514.1; --  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax toxins.  
DR Pfam; PF03495; Anthrax toxinB.  
DR PRINTS; PR01391; BINARYTOXINB.  
DR SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;  
Query Match 19.1%; Score 720; DB 2; Length 775;  
Best Local Similarity 29.0%; Pred. No. 2.5e-26;  
Matches 223; Conservative 133; Mismatches 294; Indels 120; Gaps 30;

QY 4 QENRLNESSSQGLGYFSDLNFOAPMVVTSSTGDLSPSSLEN--IPSENQVFO 61  
DB 42 QKQK---QKMDRKGLLGYFKGDF-SNLTWFAPTEDNTLYDQGTANKLDDKQKQYQ 97  
QY 62 SATWSGFIKVKSDHYTFATPSADNHVTMVDQEVINKASNNKIRLEKGRLYQIKQYQ 121  
DB 98 SIRWIGLIQSKGKGDFTNLSEDEQAIIEIDGKIIISNKGKEQVHLEKEKLVPIKIEYQ 157  
QY 122 RENPTEKGLD-----PKLWYDSQNKKEVISSDNLQLPDLKQKSN----- 162  
DB 158 SD--TKFNIDSKTFKEPKLFKIDSQSQKQKDELRNPEFNKESREFLAKSKTNPFM 215  
QY 163 SRKKRSTAGPTVPRDNDGIPDSLEVEGYTVDVQKNTFLSPWISNIHEKGLTKYKS 222

216 QRMKRDIDED---TDTGDSIPDLWEENGYT--IQNK--VAVKW-DDKFAQGYKYKLS 267  
223 PEKWSFASDPYDPEKVTGRIDKKNVSPPEARHPLVAAPYPIVHVDVMEIILSKNEEDOSTONT 282  
268 PQAHVTGVPDYDWEKAAGDIPKSAATRNPLVAAPFPIVDMERKMLSKDNL----- 323  
283 DSETRTISKNTSTSTHTSEVHNAEVAHSPFDIGGVSAGFSNSNS---TVAIDHSL 339  
324 -----NSAEHSNNSYTYANSE-----GASIEAGFGPKGFSFGVSANYQHTET 366  
340 LAGERTW---AETMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIK 395  
367 VGSQ--WGNKSKNTEGFNSASAGYLNANVYNNVGTGGIYDAQPTTSFIL-QDSIATIT 423  
396 ADENQISQILAPNNYPSKNLAPIALNAOKDASSTPIITMYNQFLEKTKQLRLDQV 455  
424 AKSNATALSTPSGDRYPASK-EGISLKTWDDFNPHITLTKPQLDAVLNNEVIKINTDQT 482  
456 YGNIAFYFENGVRVVDGNSWSEVLPOIQTETARII--FNGKDLNVERIAAVNPSDPLE 515  
483 DGRYGIIGVDG---KAEIGDRWSPIIDEIKGTASIIIDPADGKALETHIAAKDYKNPED 539  
516 TTRPDMTLKEALKIARF--FNEPNNGNLOVQKDIIT--EFDNF-----FDOOTSON 561  
540 KT-PSLTIKEGIKIAYPESISEDKGILFVEYKNDGKVTKKQLSEENIMPYLDDETSKE 598  
562 IKNLAEELNATNYVLDKIKLNAKMMILLIRKPHYDRNNIAVGADESIVKAEHREVIN 621  
599 FERQLSDGSAKGY---DKLTPKKNITIR-----LATVLGDDDFDSAYPHENATW 647  
622 SSTEGLL-----LNIDKDIRK-----ILSGYIVE--IEDTEGLKEYI---NDR 659  
648 SDKEGNLRGLSLAIPOESKYTIPOKVKVKNYDYLITGYIKHDFTTDNESIGIVAF--KKON 707  
660 YDMLNITSSLRQDGKTFIDFKYNDKPLVYSNPYKVVYANTKENTIN 709  
708 FEMWNGTISFQSGGFEKFKTKTQ-NISG-DYILDSIQLMKRNDVN 755

RESULT 9  
Q8KYK2  
ID Q8KYK2 PRELIMINARY; PRT; 225 AA.  
AC Q8KYK2;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Protective antigen-related protein, (pXO1-111).  
GN BX10133.  
OS Bacillus anthracis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STEAIN=A2012;  
RX MEDLINE=22061436; PubMed=12004073;  
RA Read T.D., Salzberg S.H., Pop M., Shumway M., Umayam L., Jiang L.,  
RA Holtzapflee E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,  
RA Keim P., Fraser C.M.;  
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in  
RT Bacillus anthracis.";  
RL Science 296:2028-2033 (2002).  
DR EMBL; AE011130; AAM26108.1; --  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
KW Plasmid.  
SQ SEQUENCE 225 AA; 25402 MW; 2E121BES4295F9C8 CRC64;  
Query Match 7.9%; Score 296.5; DB 2; Length 225;  
Best Local Similarity 37.2%; Pred. No. 6.4e-07;  
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;

QY 566 LAELNATNYVLDKIKLNAKMMILLIRKPHYDRNNIAVGADESIVKAEHREVINSTE 625

```
Db 1 MESLGINNIYNALDRKLNKNNILVRDP-YHYDNGNIVGVDDSYLKNAYKQILNWSDD 59
Qy 626 GLLNIDKIDKILSGYIVEIEDTE-----GLKEVINDYDMLNISLRODQ 672
Db 60 GVSUMLDEWQALSGYMLQKKPSNHLTNSPVITLAGKDSGVGELYRVLS-----DG 113
Qy 673 KTFIDFKYNDKPLIYISPNKYKVYAVTKEN-TIINPSNGPTSTNGIKKILIFSKKG 731
Db 114 TGFIDFNKFDENRSLV-DPGDDVYVAVTKEDFNNAVTRDENGIA-NKLKNTLVLSGKI 171
Qy 732 YEI 734
Db 172 KEI 174

RESULT 10
Q8RGK2 PRELIMINARY; PRT; 1881 AA.
AC Q8RGK2;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DE Hemolysin.
DE Hemolysin.
GN FN0291.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21866394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen M., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyripides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AEO10541; AAL94497.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008638; Haemagg.act.
DR InterPro; IPR000508; Peptidase_S26.
DR InterPro; IPR001452; SH3.
DR Pfam; PF05860; Haemagg.act; 1.
DR Pfam; PF00018; SH3; 1.
DR PROSITE; PS00761; SPASE_I_3; 1.
KW Complete proteome.
SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;

Query Match 5.9%; Score 222.5; DB 16; Length 1881;
Best Local Similarity 20.5%; Pred. No. 0.034;
Matches 173; Conservative 133; Mismatches 303; Indels 233; Gaps 36;

Qy 30 QAPMVVTS-----STTG-----DLSIPSSGELE---NIPSENV-FOSAIWSGFI 69
Db 246 QAPVATKGDVVISKGVYLDQTQAKRDIKSSSTETEGSKLLAENAINIKSGKTSNG 305
Qy 70 KVKSDYFTATADNHVTWVDDQEVN-----KASNSKILEKGR 112
Db 306 QIRANNITINGNVDSNLIFTNKDITISGLKNSGVSSSNLNVKEIENSKVYVEE-K 364
Qy 113 LYGIKI-----QYQRENPEKGLDFKLYWT--DSQNKKEVISDNLQLPELKQKSS 161
Db 365 LSSTKITNLNLSAKIEKINIFNSGKILFSKNITAKDFKNGEV-SSENLITTNLENSK 423
Qy 162 -----NSRKGSSTAGTPVPRDNDG----- 182
Db 424 INVKENINSIVNKTNAETTSKNTLNLTNRGNTIINNVSSGVIANNGKLLVGNITN 483
```

```
Qy 183 ---IPDSLEVEGYTVDKRNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKV 239
Db 484 SQNLNTATVQCKTLTDIKKINSKGLSDNLTLDIPSSGNISAKVITTOELINSGEII 543
Qy 240 TGRIDKNVSPERHPLVAAPYPIVHVDMEN-----ILLSKQEDQSTQNTDSTRISKNT 293
Db 544 SNNLSNNNNNSKIFVNGNLKISNNLNSGVIEGLELNTSIENGTITKNKLTSQL 603
Qy 294 STSRTHTSVHGNAEVHASFDFIGGSVSAFSGNSSTVA---IDHSLSLAGERWAEFM 350
Db 604 NNKK-----NTANVNAAGFLDVHNKISS-VGNIKAITMKTNNLNSGNIL-----IN 648
Qy 351 GLNTAD-----TARLNANIRYVNTGTAPIYVNLPTLSVLGKQOTLATIKADENQLSQ 403
Db 649 SLTTAENINKSGITAKNISQNLVNSGS-----VISDNITVA-----DN 687
Qy 404 ILAPNNYPSKNLAPIALNAQKASSTPITMNYNOFLELEKTKQLRLDQDVYGNIAATN 463
Db 688 ITNNTNIPANEK-----ISADKISNSK-----LVAKNTEITKLTND---GNIVVK- 730
Qy 464 FENGRRVDTGNSWSEV-----LPQOETARIIFNGKOLNIVERRIAAVNPSDPLE 515
Db 731 -ENLKAKDITNSNTIKVGENLNTDKLQNSKTLIAKNINIEKSLNNGINGKITSLNANITS 789
Qy 516 TTKPDWTLKALKIARFNEPENGNLQYQCKDITEPDFENFDQQTSONIKQLAELNATNY 575
Db 790 DIKNNNGIIQAIK-----NINIKTSNDLKDGKYTANDSLNINAKSLNNGN--- 836
Qy 576 TVLD---KIKLNKAMNIIIRDKRFHYDRNIAV-----GADESVVYKEAHREVINS--T 624
Db 837 --LENDGKIKFNLGTGNLNNKISSSSNLTANETSNNGVNSIIGSEANLTITANSKLN 894
Qy 625 EGLLL-----NI-DKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNITSSLRQ 670
Db 895 EGNLLFGEIENKLTGNTITNGVTSISLGKUKIEAKOVVNDKHIISDNDLIDVNSITN 954
Qy 671 DGTFTF-----IDFKK--YNDKLLPLIYISPNKYKVYAVTKENTINPSENGDTSTNGIK 722
Db 955 KGLLYSTNNMKVDFKENFLNDKAEIYSSG-----DITIN-SENG-TFTNRVG 999
Qy 723 KI 724
Db 1000 DI 1001

RESULT 11
Q8II04 PRELIMINARY; PRT; 3468 AA.
AC Q8II04;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pfl_0371.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung B., White O., Berriman M., Hyman R.W., K.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Arguioli S.,
RA Pettea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C., R.W.,
RA Venter C.M., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
```

DR EMBL; AE014841; AAN35955.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 3468 AA; 411913 MW; 506F7D62999BA7B1 CRC64;  
  
Query Match 5.7%; Score 215.5; DB 5; Length 3468;  
Best Local Similarity 19.4%; Pred. No. 0.16; Indels 275; Gaps 40;  
Matches 162; Conservative 134; Mismatches 266;  
  
QY 55 SENQYFQSAIGWFGFKVKKSDYETIFATSDAHVMTWVDDQEV---INKASNSKIRLEKG 111  
DB 160 SENY-----TKNDKYNNSINSY-----NKEIEFLKNAIEHSLKTEKN 202  
  
QY 112 RLYQIK-----IOYRENPTEKGLDFK---LYW-----TDSQ-----KKEVISDNL 151  
DB 203 INLLKCDSDNNDLFEKKNVHLNNDYSEDLFHDNRVNQNNLKDIIKDYIHDNDI 262  
  
QY 152 QLPULKQKSSNRKRSKRSAGTVPDRDNDGIPDSLEVEGYTVDV-KNKRFFLSPWISNI 210  
DB 263 MI-----ESNDNRNDKNCFKNTNYPNEHIIIVDKTKNGNNGNNSNNFDS--ISNI 315  
  
QY 211 HEKGLTKYKSPKWSASDP---YSDPE-KVTGRIDKNVSPKARHPLVAAPIVHV- 264  
DB 316 NKK-----ISYPINNYNSSEDKTLNQINLDMSLSDSLKXNAYSFYSL 360  
  
QY 265 ---DMENILSKNEDOSTQNTD--SETRTISKNTSTRTSTSEVHGNAEVAH---SFFDIG 317  
DB 361 EXNNNNIGMHHKNNWEYRVIDEDINISKFSNISLHSEKNDNNEFNSLCSP--- 416  
  
QY 318 GSVSAGFNSNSSTVAIDHSLSLAGERTWAETMGL-----NTADTARLANIRY 366  
DB 417 -----NSSKCKVDEMITQVGNRKVYKTMDMANDNMKNMTECHSINMESNHFKNQ 468  
  
QY 367 -VNTGTAPIYVLPVTSVLGKQ---TLATIKADE--NOLSOILAPNN--- 409  
DB 469 KINSKEDANNLSSHLNANKKIGQVNSLSDYLSIKQESNMNLSNNEALNINV 528  
  
QY 410 -----YYPKKNLAPIALNAQKASSTPITVNYNQ---FLELE 443  
DB 529 TNNNGSSNNKHSNVYKTSQVYVNPNDIN--NNHQLNLSYKNSNNLNTSNGFKKIP 586  
  
QY 444 KTKQL--RLDTQ-----VYGN-----ATYVNFGRVVDVTGNSWSEVLQ 483  
DB 587 KKNKIIISNIDFNNIFKSVIKENVKNQESNINHOFVEKNTYNDSEINKN-----NNIENN 642  
  
QY 484 IQETAR-----IIFGKDLNLVERRIAANVPSPDPLETTK----- 518  
DB 643 TQNTTCNNFINTDDIINKKIKKIEYKIDISISLLNDLSLNKLESINSIMDRYTKNVE 702  
  
QY 519 -----PDMTLKEAL-----KIAFGNEPGLQYQKDIETFDNFDDQ 557  
DB 703 EKFLDDVILDDSI: FATSNELOHSNYTTTTHIFDNNNNNNNNINOKEDLFQNDYN---K 759  
  
QY 558 TSONIKNQLAELNATNYTVDIKLNKAKNILLIRKRFHYDRNIIAVGADSESVVKEARH 617  
DB 760 ETYNNIMLSENNAELF-----KISYSCNDLVLGKNEIILD--R 798  
  
QY 618 EVINSSTEGLLANIDKIRKILSGYIEVEDTEGLKE-----VINDRYDMLNISS-- 667  
DB 799 NVENSKE-----QVEEYQNKEDIKLYHKDDNIIIDNNYDNIKKDC 842  
  
QY 668 -LRQDGKTFDFKYYND---KLPL-----YISNPNYKN-VYAVTKENTINPSEN 713  
DB 843 HLKMDNQDNINKKSDSKQKKNFINNECNVHKNFKINEIDHFEKENTLVYESKN 899  
  
PRELIMINARY; PRT; 4688 AA.  
  
RESULT 12  
Q9PQ08  
ID Q9PQ08  
AC Q9PQ08  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DR 01-OCT-2000 (TrEMBLrel. 15, last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, last annotation update)  
DE Hypothetical protein UU482.

GN UU482.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]\_SEQUENCE FROM N.A.  
RP STRAIN=Serovar 3;  
RC MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum";  
RL Nature 407:757-762 (2000).  
DR EMBL; AR02145; AAF30894.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 4688 AA; 534880 MW; B53ABFAFEZ1997E CRC64;  
  
Query Match 5.4%; Score 202; DB 16; Length 4688;  
Best Local Similarity 21.3%; Pred. No. 1;  
Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;  
  
QY 19 LLGYVESDLNFQAPMVVTSITGDLSPSSSELENIPSENQYFQSAIWSGFIKVKKSDEYT 78  
DB 3699 LVDVYYLD-NIHNQIDETRKIFKDNHV-SKEIINPGVTMISKHGNWKSPTDTTANFEFK 3756  
  
QY 79 FATSADNHVTMVYDQEVINKASNSKIRLEKRLYQIKIYQORENPTEKGLDFKLYWTD 138  
DB 3757 IETQ-----DDNDVLNNIDATVKFKDEHNNIKQKIVRIKEN-----ND 3795  
  
QY 139 SQNKKEVISDNILQELKQKSN-----SRKRSISAGTVPDRD----- 179  
DB 3796 WLKGGI---DNLN-PETKYKLENIELSPLKTHNLNLSINDKENISLITGTGPNLVK 3851  
  
QY 180 ---NGQIPDSLEVEGYTVDVKNKRTFLSPWISNTHKGLTKYKSSPEKWSATSDPYSD 235  
DB 3852 IQQNTTINDTQQTINVTLSGVNSK-YNGRKIVKYKDNNNVYESS---LITLQKQKD 3907  
  
QY 236 FEKVTGRIDKNVSPKARHPLVAAPIVHVDMENIILSKN-EDQSTQNTDSEI-RISKNT 293  
DB 3908 YQLLSNLNSN-----REYRFEKIEINHSINTNPFEDLEKUNGVSNTFITQTKT 3957  
  
QY 294 STSRTHS-EVHGNAEVHASF-----FDIGGSVSAGFS-----NSNS---STVAID 335  
DB 3958 TVQWNUSSATIVTGRVGNFNFKIKSEDKILENNQVAVFAPKETIRDTNLTWLYREPLK 4017  
  
QY 336 HSLSLAGERWTAETMGLNT---ADTARLANIRYVGTAPIYVLPVTSVLGKQNTLA 392  
DB 4018 DVTSDPKEGTWADLSNSVNFKEETTYKLVKIQVNRKPKAKKNINNSNNVILDTNISI 4077  
  
QY 393 ---TIKADENOLSOILAPNNVYPSKNLAPALN-AOKDASSTPITWNY--NQFLELE 443  
DB 4078 NSNYETTKVGDHKLINITSNNVNTNSQTINFTLSGVKSVGKKIKLSYKSNDSSEI 4137  
  
QY 444 KTKQLRLDTPQYVGNIAIYFNGR-----VRVDTGNSNWSEVLQIQETTABIIENGKD 497  
DB 4138 HTNEVLIESNKTQYINLLNLLKRNRTYTLIDVKLIDNNNVSDPFKEGNLTNSPITRTSA 4197  
  
QY 498 LNLVERIAAVNPSDPLETT-----KPDMTLKEA-----LKIAFGNEPGLNQ 541  
DB 4198 INVNLBEISNRABTNLKSIIILNLDNPNVLRDQOATIVYGNKQKQNGFITVSGNIK 4257  
  
QY 542 YQGGKDIETFDNFDDQTSQNIK-NQLAELNATNYTVDIKLNKAKNIIILIRKRFHYD- 599  
DB 4258 YLTATLDLNFN-DKVNIVNISFNKPSIAAEN-----IGIDKSNII-----YVND 4304  
  
QY 600 -----RNIIAAGA---DESUVKEAHEVINSSTEGLLANIDKIRKILSGYIEVEDTE 650  
DB 4305 IPKLEINNDIIVNGPINKKEIVVKNANOK--NNIDVDLGLQINPKIAHLR-FIAKFKSTN 4361  
  
QY 651 GLKEVIND--RYDMLNISSL-RODGKTFIDF-----KKY----- 681  
DB 4362 -----NDIIEITVINGSSLVNNDGKTSIRFTLNNLKANKLYSLVDVYLVNNSNTIVE 4415

```
QY 682 NDKPLVYISPNYKVNVAVTYKENTINPSENGDTS 717
DB 4416 SNKLP-KLNNINVOIK---INKSGHTII--SKNGEWS 4445

RESULT 13
QBIHY4 PRELIMINARY; PRT; 2849 AA.
AC QBIHY4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pf11.0392.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014841; AAN35975.1; --
KW Hypothetical protein.
SQ SEQUENCE 2849 AA; 335916 MW; B5515D173D96B813 CRC64;

Query Match 5.3%; Score 200.5; DB 5; Length 2849;
Best Local Similarity 20.6%; Pred. No. 0.64;
Matches 170; Conservative 133; Mismatches 334; Indels 189; Gaps 41;

QY 3 KOENRLNESSSQGLGY-YFDLAFQ--APMVVTSITGDLSP--SSELENIPSEN 57
DB 1862 KHNYTFNKEENLKSIFKYNNNNNINEKDDIPKSIQSSFNEDNIAVYKNGNMKLKN 1921
QY 58 Q-YFQSAIWSGFIKVKKDEYTFATSDADNHVTMWDD-----QEVINKASNSK--- 105
DB 1922 EDIQRFSHNKIKYENKNCSDSCNIVKINYDELNDSTQTKELNEGKSNNNGKAEAW 1981
QY 106 -IRLEKGRLY-QIKIQYORENTEKGLDFKLYWTDSONKKEVISEDNLQLPKQKSSNS 163
DB 1982 IIDIKNNETYPYKIDKKEKNEK--KNKYMNKNNDNKNMKGSSN---KSMCKNNKNS 2036
QY 164 RKRSTSGACTVPDRDNGIDPDSLEVGTVVDKVKRTFLSPWISNTHEKGLTKYKSSP 223
DB 2037 NKQKH-----IPLSVNNKY-----NKSSINKYENNIN-----KNK 2069
QY 224 EKWSTASDPYDF--EKVTGRI-DKNVSPERHPLVAAYPIVHDM-----NILS 272
DB 2070 DKLNLVNSISKLQVSKIKQELSNKISKD-----ILNFEITKIKKSKKETKNTNTN 2124
QY 273 KXEDOSTQNTDSETTISQNTSTSTHTSEV-HGNAEVHAFPDIGGSVAG-PSNSNSS 330
DB 2125 KNDNNNDNNINFRINDNINSYVNLANKVPYDINI-----DEGNVINTGPIYADGS 2179
QY 331 TVAIDHLSLAGERTWAETWGLNTA-----DTARLNA-----NIRYVNTGTAPIYVNL 378
DB 2180 TI-----YTWANNIDTNYMYNKYFDSKKNINHVPLLNVPYLN-----NDL 2221
QY 379 PTTSLVLGKQIATIKADENQLSQ-----ILAPNYYPSKLAIPALNAQKASSTPTM 434
DB 2222 LINNVILNQNMNNLENLNTNTIGSVQPEVTCPDFY--ANNIKSYLD-----PNLQ 2271

QY 435 NYNOFLELEKTKQLRLDT-----DOVYGNATYTFENGREVDT-----GSN 476
DB 2272 NNNYFDNNQLLNHNLANNLNSINYIDNYLSYNNLCINGINIKOTCKDIVIGIPNTN 2331
QY 477 WSEVLPOIQETTARIIFNGKDLNLVERRIAAVNPSPDELE-----TTKPDMTLKALK 528
DB 2332 QNOI-PTI-ELDDTILKN--DVNLLNNNNVNVHNSVANVEMLNINQNVQKLYNDIQENLH 2387
QY 529 IAFGF--NEPNGNLOYQKGD-----ITEFPNFDQOTSQNIKNQLAELNATNIYTV 577
DB 2388 ITNSLHNNNNNNNNMLGGDVLLNNSYLFNINSFPNNMAYIYNNNNNNNNNNCINYL 2447
QY 578 LDKIKLNK-----WNILIRDKRPHYDRNNIAVGADESVVKEAHEVINSSTEGLLNID 632
DB 2448 LDKKEINVKEINEINWNSVLEDTNKEGND-----MKDNKYIINK--ENNVMKND 2496
QY 633 KDIRKILSGYVIEIDTEGLKEVINDRYDMLNISLRODGTFFDKYNDKLPYISNP 692
DB 2497 NNIDIISNGQINNTNANTMKERNLNRDDDIYKKNMKRNSLDFKGLDSE-----KK 2550
QY 693 NYKVVV---YAVTKENT--IINPSENGDTSTNGIKKILIFSKGYE 733
DB 2551 NKSMNIEGARKNKINTTQLLNYSEN---RKGLRDMXYADKVL 2593

RESULT 14
QBIKS2 PRELIMINARY; PRT; 1904 AA.
AC QBIKS2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Ferlin, putative.
GN Pf14.0530.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Perte M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014825; AAN37143.1; --
DR InterPro; IPR000008; C2.
DR Pfam; PF00168; C2; 5.
DR SMART; SM00239; C2; 5.
DR PROSITE; PS00004; C2_DOMAIN_2; 2. A36470367BDEC41 CRC64;
SQ SEQUENCE 1904 AA; 224720 MW; A36470367BDEC41 CRC64;

Query Match 5.3%; Score 198.5; DB 5; Length 1904;
Best Local Similarity 19.8%; Pred. No. 0.47;
Matches 174; Conservative 128; Mismatches 328; Indels 247; Gaps 43;

QY 3 KOENRLNESSSQGLGYYSFDLN-FQAPMVVTSITGDLSPSSELENIPSENQYFQ 61
DB 836 KQNN--NKQNNNNNNNNYNNSSVYQNDLLYGERVGDILL---YFELVQSKA-MK 889
QY 62 SAJWSGFIKVKKDEYTFATSDADNHVTMWDDQEVINKASNSKIKLEKGRLYQIK--- 117
DB 890 FPIYPMITEIKTKTLTFFCNSLENILM-----KXANFLKLSFRNNKYQISTPII 941
QY 118 -----IQYQRENTKGLDFKLYWTDSONKKE--EVLSSDNLQLPK 157
```



Db 942 LLSITSYSSYGGKKXVLMKYEKTLKANTRIQLK-AWKNFNOQSFEMFIENNNI----- 996  
Qy 158 QKSSNRKRKSTAGPTVPRDNDGIPDS-LEVEGYTVDVKNK-----RTFLS-----PWI 207  
Db 997 -----DIPLOPIFDPIILNIKVKYKVKYFGETNISLVPLPWI 1037  
Qy 208 SNIHEKKGLTKYKSSPEKWTASDPYDFPKV-----TGRIDKXVSPPEARHPLVAAYPIV 262  
Db 1038 KNIDE-----VLY-----YLOAHDYSETINMKNDINTFYKN-----KXAAALVTSATS 1082  
Qy 263 HVDMENILSKNEDQSTQNTDSET-----RTISKNTSTSRHTSEVHGNAEVAHAFSD 315  
Db 1083 LADCEDTSLKEELNKYENDDDDEAWKEIPLNLDQENQKEDKNKTSCHGVNNTYDGYN 1142  
Qy 316 IGGVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTARKLANIRYVNTGTAPI- 374  
Db 1143 -NGAYEMGYMETYNIKNDN-----NNNNNNYNNNNNNNNNNNNNNNNNNNNYNYAAPT 1192  
Qy 375 ---NVNLPFTSLVGLGKNTLATIKAD--ENOLSOILAPNNYVPSKILAPIALNAQDASS 429  
Db 1193 SYNINV-----QNDTRNNRYNHSNNMMNNNNNNNNNNNNNNNNNNNNNNNNNN 1244  
Qy 430 TPIITWYN-----OFLELEKTK-----QLRLTDQVYGNIAATN--PENGRRVVDGTGSNW 477  
Db 1245 KGTILFNNNNNHHFNKLSNNKFSYLSRIQKDTYNIKNYNNIYKLFDDGTPETIILSYN 1304  
Qy 478 SEVLPOIQTETARIEN-----GKDLNLVERRIAAYN----- 509  
Db 1305 VANYPIKILSKYILNHIIPRFLYVGGKLN--IEKFIKNIRVSVGILENYLDDIL 1363  
Qy 510 -PSDPLETTKPDMTL-----KEALKIA-----FGFNEPNG-----NLQYQCKDI 547  
Db 1364 IPSLPLIKKNDISCDNNYNNENKIEKQKFCQFPFVEIIGGQIKCFKIKYRNLES 1423  
Qy 548 TEFDENFQOQSONI-----KNOLA-----ELNATNIYVLOKIKLNKAGNILIR 592  
Db 1424 ENNPUSLKDITNQIFRNKFRGNKIPYLKIRVYVLRGILYGINNEYTANPLYIFSLG 1483  
Qy 593 DK-----RFRYDRNNI-----AVGDESIVYKHAHREVINSSTEGLLNIDKIRKILSGYI 643  
Db 1484 EKTSLNRAFKFSNINPEFGCLWESEAFPE--DEILTSIVSAEDNYDKQINDIYIG-- 1539  
Qy 644 VEIETEGLEKEVIN--DRYDMLNITSLRQDGTDFDKK--YND--KLPLYISNPNYKV-- 696  
Db 1540 -----STE-----INLFRWMSKEWRHMKKPKIPVEYRPLYSNIXIKHPKMYVSSNNYTMN 1590  
Qy 697 ---NVYAV-----TKENTIIINPSENGDTS 717  
Db 1591 SWNNIFSPFDIPNLYMTVTSPTKGNNNNNNNNNNNNN 1627  
Q8IHP9 PRELIMINARY; PRT; 2940 AA.  
AC Q8IHP9;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN Pfl1\_0480.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]\_TaxID=36329;  
RP SEQUENCE FROM N.A.  
RC STRAIN=3D7;  
RX MEDLINE=22255705; PubMed=1236864;  
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,  
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,

RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
RT Fraser C.M., Barrell B.,  
RT "Genome sequence of the human malaria parasite Plasmodium  
RT falciparum";  
RL Nature 419:498-511 (2002).  
DR EMBL; AE014843; AAN36060.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 2940 AA; 348040 MM; BF8748D8A051BAD2 CRC64;  
Query Match 5.2%; Score 197.5; DB 5; Length 2940;  
Best Local Similarity 18.7%; Pred. No. 0.92;  
Matches 162; Conservative 138; Mismatches 322; Indels 243; Gaps 40;  
Qy 34 VVTSSTGDLSTPSSSELENIPSENCYFQSAIWSGFIKVKYSDEYTFATSAADHVTMWDD 93  
Db 627 ITHNNNDNISINNNNNNN-----NYIFMNNYMHNNI-----NNNYCYCYNNNTNHNYYNN 679  
Qy 94 QEVINKASNSK-----IRLEKGRLYQIKIOYQRENPTTEKGLDFK-- 133  
Db 680 IYIQNHQDQNPAPILQPINNHIAHINDLCYHSEKNEYTKISKHOMNINPQOSNGKNQ 739  
Qy 134 -----LYWTDSONK-----KEVISDNLQLPKQKSSNRKSTSGPT-- 174  
Db 740 NDISNNINKNDYYNLNEQKILCDKXSYIKCDIPQKCDNTQDDENSEQONQYITNPSG 799  
Qy 175 -----VPRDNDGIP--DSLEVEGYTVDVKNKRTFLSPWISNIEKKGK-- 216  
Db 800 HYKIKEQMYVQHIPIDEHDTNTEMINQNTY-----NINLDPYIMNNDNNVYLNQ 853  
Qy 217 ----TKYKSSPE--KWTASDPYSB----FEKVTGRIDK--NVSPARPLVAAPIVHVD 265  
Db 854 NYFDEKCKKKEEDIMITEANTHNNVHTYQNNSHNMKNMLNNNTNSKSTQDYDLYHNN 913  
Qy 266 ME-----NITLSK--NEDQST-----QNTDSTRTAISKNTS--T 295  
Db 914 MENFNTYNNIKKEKIHNDTSSVNNSSINQQNLCKNKNTNDYQNRINKDNSIPD 973  
Qy 296 SPTHSEVHGNAEVAHAFSDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETWGLNTA 355  
Db 974 SSINNELKNNIQMDIY--ASNTYNNNNNNNNNNNNFIISNNI--ISYNNY--MNES 1024  
Qy 356 DT--ARLAN--TRYVNTGTAPYVNLVLP--TTSVLGKNTLATIKADENQLS-----Q 403  
Db 1025 NVTPQNSNNYPIYD-----PHLNPEYKHTINNQLNKNKNNNNISMDNVYDSNTISH 1080  
Qy 404 ILAPNNYPSKNLAPIALNAQDASSTPIITMNYNQFLELEKTKQLRLTDQVYGNIAATN 463  
Db 1081 ISVNDNEYNSNFTNDMYNNNN-----NNNNNNNNNNNNNNNNNNNNNNNNNNNN 1132  
Qy 464 FENGRRVVDGTGSNWSEVLPCIOETTA--RIIFNG--KDLNLVERRIAAYNPSDPLETTKPD 521  
Db 1133 LYNSTNYITQHPSSNDVFKIOKHANNIIMGHKEEHLIEKK-----SKEETNKYNE 1185  
Qy 522 TLKEALKIAFGFNEPNGNLQYQCKDITFDFNFDOQTSONIKNQLAEL-----NA 571  
Db 1186 QVYRSI-----NQNTLILKENEIDENDINTLQ--NLNKKDMMNDNINLKNLANI 1237  
Qy 572 TNYITVLDKILNAKQNIILIRKRFHYDRANN--AVGADESIVYKHAHREVINSSTEGLLNI 631  
Db 1238 NNIYT-----PYQNILKNNIEQFLNN-----KEVITKHAYTHSSNEININVKNI 1284  
Qy 632 DKDIRKILSGYIVEIEDTEGLEKEVINDRYDMLNITSLRQDGTK-----ID- 677  
Db 1285 DTQI-----NIRKNNNDIILNK--EQKDISTNNNQNTYNTVTSVKNVEYSILDN 1334  
Qy 678 ---FKYNDKPLIYINP-----NYKNVYAVTKENT-----IINPSENG 714  
Db 1335 SPTDKYVKKNLNLISSSYNTNDHNNYHDDIDDKTKNKNKYEKHKDKINYYNNINSNG 1394  
Qy 715 DTSTNGI-----KKILIF 727

Db 1395 NKQNFINSYFDLNENEKKKINIF 1419

Search completed: May 3, 2004, 19:40:06  
Job time : 37.6261 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 47,5168 seconds  
(without alignments)  
4227.791 Million cell updates/sec

Title: US-09-848-909A-16  
Perfect score: 3648  
Sequence: 1 EVKQENRLNSESSESSQGLL.....TSTNGIKKILFSSKGYEIG 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5  
Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003Bs.\*  
8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3648	100.0	711	5	AAM51498 Anthrax P
2	3642	99.8	711	5	AAM51497 Anthrax P
3	3635	99.6	711	5	AAM51496 Anthrax P
4	3629	99.5	711	5	AAM52113 Anthrax P
5	3626	99.4	735	5	AAM51495 Anthrax P
6	3620	99.2	735	5	AAM51493 Anthrax P
7	3614	99.1	735	5	AAM51485 Anthrax P
8	3613	99.0	735	5	AAM51491 Anthrax P
9	3613	99.0	735	5	AAM51492 Anthrax P
10	3603	98.9	735	5	AAM51490 Anthrax P
11	3608	98.9	735	5	AAM51487 Anthrax P
12	3608	98.9	735	5	AAM51500 Anthrax P
13	3608	98.9	735	5	AAM51489 Anthrax P
14	3607	98.9	735	2	AAR60179 Anthrax P
15	3607	98.9	735	5	AAM51483 Anthrax P
16	3607	98.9	735	5	AAM51488 Anthrax P
17	3607	98.9	735	5	AAM51494 Anthrax P
18	3607	98.9	735	5	AAM51499 Anthrax P
19	3607	98.9	736	3	AAY56959 B. anthra
20	3607	98.9	763	3	AAY56960 B. anthra
21	3607	98.9	764	3	AAY56958 B. anthra
22	3607	98.9	764	4	AAB47306 Wild type
23	3607	98.9	857	7	AD65872 Bacillus
24	3606	98.8	735	5	AAM51484 Anthrax P
25	3605	98.8	735	5	AAM51486 Anthrax P

RESULT 1  
AAM51498  
ID AAM51498 standard; protein; 711 AA.  
XX  
AC AAM51498;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Anthrax PA K397D/D425K/F427A/D2L2 deletion mutant.  
XX  
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutuin.  
XX  
OS Bacillus anthracis.  
OS Synthetic.  
FH  
FH Key Location/Qualifiers  
FT Misc-difference 302..303  
FT /note= "The mutant comprises a deletion amino acids 302-325 of the D2L2 loop of the wildtype protein"  
FT  
FT Misc-difference 373  
FT /note= "Wild-type Lys substituted by Asp"  
FT  
FT Misc-difference 401  
FT /note= "Wild-type Asp substituted by Lys"  
FT  
FT Misc-difference 403  
FT /note= "Wild-type Phe substituted by Ala"  
WO200182798-A2.  
08-NOV-2001.  
04-MAY-2001; 2001WO-US014372.  
04-MAY-2000; 2000US-0201800P.  
(HARD ) HARVARD COLLEGE.  
Collier RJ, Sellman BR;  
WPI; 2002-017725/02.  
Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.  
Claim 4; Page; 77pp; English.  
The invention relates to antibacterial agents comprising mutant forms of

ALIGNMENTS

26	3604	98.8	735	5	AAM518289
27	3604	98.8	764	6	AAM535717
28	3592	98.5	764	6	ABP71693 B. anthra
29	3562.5	97.7	903	2	AAM60183 PA(1-725)
30	3430	94.0	719	2	AAM60193 Modified
31	2891	79.2	595	5	AAM58288 Bacillus
32	2752	75.4	569	3	AAY56961 B. anthra
33	2342	64.2	487	5	AAM518285
34	2027	55.6	426	5	AAM518287
35	1478	40.5	318	5	AAM518286
36	1338	36.7	258	5	AAM518284
37	1275	35.0	288	5	AAM50707 Bacillus
38	821.5	22.5	1052	4	AAM507903
39	820	22.5	1032	4	AAM507901 C. botuli
40	816	22.4	1092	4	AAM507900 C. botuli
41	807.5	22.1	1112	4	AAM507902 C. botuli
42	792	21.7	721	6	AAM535719 Clostridi
43	759.5	20.8	880	2	AAM60224 Bacillus
44	759	20.8	881	3	AAY59277 MTS toxin
45	758	20.8	884	2	AAR91239 B. cereus

CC	pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants in the B moiety of the pore-forming binary A-B anthrax toxin, where the B moiety is anthrax protective antigen (PA) and using these mutants or combinations of them for protecting against Bacillus anthracis infections in humans, especially as vaccines. Note: The present sequence is not given in the specification but is derived from the Bacillus anthracis wild-type PA protein sequence shown in figure 13 (AAMS1483)	
XX	Sequence 711 AA;	
SY	Query Match 100.0%; Score 3648; DB 5; Length 711; Best Local Similarity 100.0%; Pred. No. 7.6e-241; Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 EVKQENRLNSESSESSQGLGYPFDLNFQAPVWVTSSTTGDLSIPSELENIPSENQYF 60	
DB	1 EVKQENRLNSESSESSQGLGYPFDLNFQAPVWVTSSTTGDLSIPSELENIPSENQYF 60	
QY	61 QSAIWSGFIKYKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120	
DB	61 QSAIWSGFIKYKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120	
QY	121 QRENTEKGLDFKLYWTDSONKKEVVISDNLQLPELKQKSSNRKCRSTAGTVPDRDN 180	
DB	121 QRENTEKGLDFKLYWTDSONKKEVVISDNLQLPELKQKSSNRKCRSTAGTVPDRDN 180	
QY	181 DGIIPDSLEVGTVYDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASTADSPYDFEKT 240	
DB	181 DGIIPDSLEVGTVYDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASTADSPYDFEKT 240	
QY	241 GRIDKNVSPEARHPLVAAPYVHVVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300	
DB	241 GRIDKNVSPEARHPLVAAPYVHVVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300	
QY	301 SNSNSTVAIDHSLAGERTWAETMGLNTADTARLNNANIRYNTGTAPIYVNLPTTSIV 360	
DB	301 SNSNSTVAIDHSLAGERTWAETMGLNTADTARLNNANIRYNTGTAPIYVNLPTTSIV 360	
QY	361 LGKNQPLATIKADENQLSQILAPNNYPSKNLAPIALNAQKADASTPTIMYNNQFLELEK 420	
DB	361 LGKNQPLATIKADENQLSQILAPNNYPSKNLAPIALNAQKADASTPTIMYNNQFLELEK 420	
QY	421 TQQLRLDTPQVYGNATYTNFNGRVRVDTGSNWSEVLPOIQTETARIIFNGKDLNVER 480	
DB	421 TQQLRLDTPQVYGNATYTNFNGRVRVDTGSNWSEVLPOIQTETARIIFNGKDLNVER 480	
QY	481 IAAVNSPDPLETTKPDWTLKEALKIAGFNPNPNGLQYQKGKITEFDNFDQOTSQINX 540	
DB	481 IAAVNSPDPLETTKPDWTLKEALKIAGFNPNPNGLQYQKGKITEFDNFDQOTSQINX 540	
QY	541 QLAELNATNIYVLDKIKLNKQNTLIEDKSPHYDRNNIAGVADSVVKEAHREVINSST 600	
DB	541 QLAELNATNIYVLDKIKLNKQNTLIEDKSPHYDRNNIAGVADSVVKEAHREVINSST 600	
QY	601 EGGLELIDDKIRKILSGYIVEIDTEGLKEVINDRYDMLNISSLRQDGTTFIDPKKYNK 660	
DB	601 EGGLELIDDKIRKILSGYIVEIDTEGLKEVINDRYDMLNISSLRQDGTTFIDPKKYNK 660	
QY	661 LPLVISPNYKNNVAVTNTIINPSENGDTSTNGIKKILLFSKKGYEIG 711	
DB	661 LPLVISPNYKNNVAVTNTIINPSENGDTSTNGIKKILLFSKKGYEIG 711	
RESULT 2		
AAMS1497		
ID	AAWS1497 standard; protein; 711 AA.	
XX	AAWS1497;	
AC	AAWS1497;	
XX	01-FEB-2002 (first entry)	
DT	Anthrax PA K397D/F427A/D212 deletion mutant.	
XX		

361 LGKQTLATIKADENQSLQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420  
361 LGKQTLATIKADENQSLQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420  
421 TKQLRLDQVGNATYFENGVRVDTGSNWSEVLPOIQETTARIIFNGKDLNLVERR 480  
421 TKQLRLDQVGNATYFENGVRVDTGSNWSEVLPOIQETTARIIFNGKDLNLVERR 480  
481 IAAVNPSPDPLETTKPDMTLKEALKIAPFNGFNEPQNGNLYQYQKDIETFDNFDDQTSQNIKN 540  
481 IAAVNPSPDPLETTKPDMTLKEALKIAPFNGFNEPQNGNLYQYQKDIETFDNFDDQTSQNIKN 540  
541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNAIAGADESVVKEAHEVINSST 600  
541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNAIAGADESVVKEAHEVINSST 600  
601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGKTFIDFKKYNK 660  
601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGKTFIDFKKYNK 660  
661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711  
661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711

RESULT 3  
AAW51496  
ID AAW51496 standard; protein; 711 AA.  
XX  
AC AAW51496;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Anthrax PA F427A/D2L2 deletion mutant.  
XX  
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX  
OS Bacillus anthracis.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 302..303  
FT /note= "The mutant comprises a deletion amino acids 302-325 of the D2L2 loop of the wildtype protein"  
FT  
FT Misc-difference 403  
FT /note= "Wild-type Phe substituted by Ala"  
XX  
PN WO200182788-A2.  
XX  
XX 08-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-05014372.  
XX  
XX 04-MAY-2000; 2000US-0201800P.  
XX  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX  
XX Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.  
XX  
XX Claim 4; Page; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of pore-forming toxins (AAW52113 and AAW51484-AAW51500), especially mutants in the B moiety of the pore-forming binary A-B anthrax toxin, where the B moiety is anthrax protective antigen (PA) and using these mutants or

CC compositions of them for protecting against Bacillus anthracis infections in humans, especially as vaccines. Note: The present sequence is not given in the specification but is derived from the Bacillus anthracis CC wild-type PA protein sequence shown in figure 13 (AAW51483)  
XX  
SQ Sequence 711 AA;  
Query Match 99.6%; Score 3635; DB 5; Length 711;  
Best Local Similarity 99.7%; Pred. No. 5.9e-240;  
Matches 709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSTTODLSIPSELENISENOYF 60  
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSTTODLSIPSELENISENOYF 60  
QY 61 QSAIWGSIKVKSDSEYTFATSDNHVTWMDQDEVINKASNKIRLEKGLYQIKIY 120  
DB 61 QSAIWGSIKVKSDSEYTFATSDNHVTWMDQDEVINKASNKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKOKSSNRKRSAGTVPDRDN 180  
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKOKSSNRKRSAGTVPDRDN 180  
QY 181 DGIPDSLEVEGYVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKWSASDPYSDFEVT 240  
DB 181 DGIPDSLEVEGYVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKWSASDPYSDFEVT 240  
QY 241 GRIDKNVSPEARHPLVAAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300  
DB 241 GRIDKNVSPEARHPLVAAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300  
QY 301 SNSNSTVAIDHSLSLAGERTWAETMGLTADTARLNANIRYVNTGTAPIYVLPPTSLV 360  
DB 301 SNSNSTVAIDHSLSLAGERTWAETMGLTADTARLNANIRYVNTGTAPIYVLPPTSLV 360  
QY 361 LGKQTLATIKADENQSLQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420  
DB 361 LGKQTLATIKADENQSLQILAPNNYPSKVLAPIALNAQKADASTPTITWYNNQFLELEK 420  
QY 421 TKQLRLDQVGNATYFENGVRVDTGSNWSEVLPOIQETTARIIFNGKDLNLVERR 480  
DB 421 TKQLRLDQVGNATYFENGVRVDTGSNWSEVLPOIQETTARIIFNGKDLNLVERR 480  
QY 481 IAAVNPSPDPLETTKPDMTLKEALKIAPFNGFNEPQNGNLYQYQKDIETFDNFDDQTSQNIKN 540  
DB 481 IAAVNPSPDPLETTKPDMTLKEALKIAPFNGFNEPQNGNLYQYQKDIETFDNFDDQTSQNIKN 540  
QY 541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNAIAGADESVVKEAHEVINSST 600  
DB 541 QLAELNATNIYTVLDKIKLNAKNILIRDRKFRHYDRNNAIAGADESVVKEAHEVINSST 600  
QY 601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGKTFIDFKKYNK 660  
DB 601 EGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNISLRQDGKTFIDFKKYNK 660  
QY 661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711  
DB 661 LPLYISNPYKVNVAVTKENTIINPSENGDTSTNGIKKILIPSKGYEIG 711

RESULT 4  
AAW52113  
ID AAW52113 standard; protein; 711 AA.  
XX  
AC AAW52113;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Anthrax PA D2L2 deletion mutant.  
XX  
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX



```

Query Match      99.4%; Score 3626; DB 5; Length 735;
Best Local Similarity 96.7%; Pred. No. 2.6e-239;
Matches 711; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120

QY 121 QRENPTKGLDPLKLYWTDQSNKKEV:SSDNQLQPLKQKSSNRKCRSTAGTVPDRDN 180
DB 121 QRENPTKGLDPLKLYWTDQSNKKEV:SSDNQLQPLKQKSSNRKCRSTAGTVPDRDN 180

QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSWTASDPYSDFEKT 240
DB 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSWTASDPYSDFEKT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRHT 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 397 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGRVVDTSNNWSEV 456
DB 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGRVVDTSNNWSEV 480

QY 457 LPOIQTETAILFNGKDLNIVERRIAANVPSDPLETTKPDMTLKALKIAPFNPNGNL 516
DB 481 LPOIQTETAILFNGKDLNIVERRIAANVPSDPLETTKPDMTLKALKIAPFNPNGNL 540

QY 517 QYQCKDITEDFDFDQTSQNIKNQLAELNATNITVLDKIKLAKMILLIRDFHVDYR 576
DB 541 QYQCKDITEDFDFDQTSQNIKNQLAELNATNITVLDKIKLAKMILLIRDFHVDYR 600

QY 577 NNIAVGADESUVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 636
DB 601 NNIAVGADESUVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660

QY 637 DMLNISLRDQDKTFIDFKYNDKFLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 696
DB 661 DMLNISLRDQDKTFIDFKYNDKFLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 697 IKKILIPSKKGYBIG 711
DB 721 IKKILIPSKKGYBIG 735

RESULT 6
ID AAM51493 standard; protein; 735 AA.
XX AAM51493;
XX 01-FEB-2002 (first entry)
XX Anthrax PA mutant K397D/D425K.
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX Bacillus anthracis.
OS Synthetic.

```

```

XX Key Location/Qualifiers
FH Misc-difference 397 /note= "Wild-type Lys substituted by Asp"
FT FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
XX W0200182788-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.
XX 04-MAY-2000; 2000US-0201800P.
XX (HARD ) HARVARD COLLEGE.
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX Sequence 735 AA;

```

```

Query Match      99.2%; Score 3620; DB 5; Length 735;
Best Local Similarity 96.6%; Pred. No. 6.6e-239;
Matches 710; Conservative 0; Mismatches 1; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120

QY 121 QRENPTKGLDPLKLYWTDQSNKKEV:SSDNQLQPLKQKSSNRKCRSTAGTVPDRDN 180
DB 121 QRENPTKGLDPLKLYWTDQSNKKEV:SSDNQLQPLKQKSSNRKCRSTAGTVPDRDN 180

QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSWTASDPYSDFEKT 240
DB 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKSWTASDPYSDFEKT 240

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRHT 300
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRHT 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396
DB 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 397 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGRVVDTSNNWSEV 456
DB 421 LNAQKDASSPTITMNNYQFLEKTKQLRLDQDVYGNIAATYNFENGRVVDTSNNWSEV 480

```

QY 457 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 516  
 DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 540  
 QY 517 QYQCKDITEFDNFDQOTSONIKQLAELNATNIYTVLDKIKLNKQNLILIRDKRFHYDR 576  
 DB 541 QYQCKDITEFDNFDQOTSONIKQLAELNATNIYTVLDKIKLNKQNLILIRDKRFHYDR 600  
 QY 577 NNIAGADESVVKEAHEVINSSTEGLLNIDKIDIKILSGYIVETEDTEGLKEVINDRY 636  
 DB 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIDIKILSGYIVETEDTEGLKEVINDRY 660  
 QY 637 DMLNSSLRQDGKTFIDFKKYNKDLPLIYISNPNKYVAVTKENTIINPSENGDTSTNG 696  
 DB 661 DMLNSSLRQDGKTFIDFKKYNKDLPLIYISNPNKYVAVTKENTIINPSENGDTSTNG 720  
 QY 697 IKKILIFSKKGYEIG 711  
 DB 721 IKKILIFSKKGYEIG 735

## RESULT 7

AAM51485  
 ID AAM51485 standard; protein; 735 AA.

AC AAM51485;

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant K397D.

QY Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

OS Bacillus anthracis.

OS Synthetic.

XX Key Location/Qualifiers

FH Misc-difference 397

FT /note= "Wild-type Lys substituted by Asp"

XX WO200182788-A2.

PD 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

PR (HARD ) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

DR Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.

PS Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)

XX Sequence 735 AA;

Query Match 99.1%; Score 3614; DB 5; Length 735;

Best Local Similarity 96.5%; Pred. No. 1.7e-236; Mismatches 2; Indels 24; Gaps 1;  
 Matches 709; Conservative 0;

QY 1 EVKQENRLLNESSESSQGLLYYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENYF 60  
 DB 1 EVKQENRLLNESSESSQGLLYYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENYF 60  
 QY 61 QSAIWGSFTKVKKSDEYTFATSDADNHVTWVDDQEVINKASNNKIRLEKGRLYQIKIY 120  
 DB 61 QSAIWGSFTKVKKSDEYTFATSDADNHVTWVDDQEVINKASNNKIRLEKGRLYQIKIY 120  
 QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSNRSKRKSTAGTPVPRDN 180  
 DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSNRSKRKSTAGTPVPRDN 180  
 QY 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASPDYDFEKT 240  
 DB 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASPDYDFEKT 240  
 QY 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRHT 300  
 DB 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRHT 300  
 QY 301 -----SNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 336  
 DB 301 SEVHGNAEVHSAFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360  
 QY 337 NANIRYVNTGTAPIYNNVLTPTSLVIGKQQTATIKADENQLSQILAPNNYPSKNLAPIA 396  
 DB 361 NANIRYVNTGTAPIYNNVLTPTSLVIGKQQTATIKADENQLSQILAPNNYPSKNLAPIA 420  
 QY 397 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 456  
 DB 421 LNAQKQDASPTPTMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480  
 QY 457 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 516  
 DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNPNGL 540  
 QY 517 QYQCKDITEFDNFDQOTSONIKQLAELNATNIYTVLDKIKLNKQNLILIRDKRFHYDR 576  
 DB 541 QYQCKDITEFDNFDQOTSONIKQLAELNATNIYTVLDKIKLNKQNLILIRDKRFHYDR 600  
 QY 577 NNIAGADESVVKEAHEVINSSTEGLLNIDKIDIKILSGYIVETEDTEGLKEVINDRY 636  
 DB 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIDIKILSGYIVETEDTEGLKEVINDRY 660  
 QY 637 DMLNSSLRQDGKTFIDFKKYNKDLPLIYISNPNKYVAVTKENTIINPSENGDTSTNG 696  
 DB 661 DMLNSSLRQDGKTFIDFKKYNKDLPLIYISNPNKYVAVTKENTIINPSENGDTSTNG 720  
 QY 697 IKKILIFSKKGYEIG 711  
 DB 721 IKKILIFSKKGYEIG 735

## RESULT 8

AAM51491  
 ID AAM51491 standard; protein; 735 AA.

AC AAM51491;

XX 01-FEB-2002 (first entry)

XX Anthrax PA mutant D425K.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
 OS Bacillus anthracis.  
 OS Synthetic.



PH Key Location/Qualifiers  
FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"  
XX W0200182788-A2.  
XX 08-NOV-2001.  
XX 04-MAY-2001; 2001WO-US014372.  
XX 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX  
XX Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX  
XX Claim 4; Page; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of  
XX pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants  
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX moiety is anthrax protective antigen (PA) and using these mutants or  
XX compositions of them for protecting against Bacillus anthracis infections  
XX in humans, especially as vaccines. Note: The present sequence is not  
XX given in the specification but is derived from the Bacillus anthracis  
XX wild-type PA protein sequence shown in figure 13 (AAMS1483)  
XX  
XX Sequence 735 AA;  
XX  
XX Query Match 99.0%; Score 3613; DB 5; Length 735;  
XX Best Local Similarity 96.5%; Pred. No. 2e-238;  
XX Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;  
XX  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSTTGDLSIPSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSTTGDLSIPSELENIPSENQYF 60  
QY 61 QSATWSGFIKKVKSDEYTFATSAADNHVTWVDDDEVINKASNSKIRLEKRLYQIKIQY 120  
DB 61 QSATWSGFIKKVKSDEYTFATSAADNHVTWVDDDEVINKASNSKIRLEKRLYQIKIQY 120  
QY 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKGSTSGPTVPDRDN 180  
DB 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKGSTSGPTVPDRDN 180  
QY 181 DCIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSFASDPYDFEKT 240  
DB 181 DCIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSFASDPYDFEKT 240  
QY 241 GRIDKNVSPERARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPERARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAGTNGINTADTARL 336  
DB 301 SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTNGINTADTARL 360  
QY 337 NANIRYVNTGTAPIVNLPTTSLVLGKQTLATIKADENQSLQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIVNLPTTSLVLGKQTLATIKADENQSLQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASTPTIMYNOFLEKTKQLRLDPTDQVYGNIAIYFNFGVRVDTGNSWSEV 456  
DB 421 LNAQKDFSTPTIMYNOFLEKTKQLRLDPTDQVYGNIAIYFNFGVRVDTGNSWSEV 480  
QY 457 LQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKALKIAFGFNEPNGNL 516

DB 481 LQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDWTLKALKIAFGFNEPNGNL 540  
QY 517 QYQKDIITFDFNFQDQTSQNIKNQALAEINATNIYTVLDIKLNAKNILIRDKRFHYDR 576  
DB 541 QYQKDIITFDFNFQDQTSQNIKNQALAEINATNIYTVLDIKLNAKNILIRDKRFHYDR 600  
QY 577 NNIAVGADESUVKGAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 636  
DB 601 NNIAVGADESUVKGAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660  
QY 637 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVVAVTKENTINPSENGDTSTNG 696  
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVVAVTKENTINPSENGDTSTNG 720  
QY 697 IKKILIFSKKGVEIG 711  
DB 721 IKKILIFSKKGVEIG 735  
XX  
XX RESULT 9  
XX AAMS1492  
XX ID AAMS1492 standard; protein; 735 AA.  
XX AC AAMS1492;  
XX DT 01-FEB-2002 (first entry)  
XX DE Anthrax PA mutant F427A.  
XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.  
XX OS Bacillus anthracis.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Misc-difference 427 /note= "Wild-type Phe substituted by Ala"  
XX PN W0200182788-A2.  
XX PD 08-NOV-2001.  
XX PF 04-MAY-2001; 2001WO-US014372.  
XX PR 04-MAY-2000; 2000US-0201800P.  
XX PA (HARD ) HARVARD COLLEGE.  
XX PI Collier RJ, Sellman BR;  
XX PS WPI; 2002-017725/02.  
XX DR Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX PS Claim 4; Page; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of  
XX pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants  
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX moiety is anthrax protective antigen (PA) and using these mutants or  
XX compositions of them for protecting against Bacillus anthracis infections  
XX in humans, especially as vaccines. Note: The present sequence is not  
XX given in the specification but is derived from the Bacillus anthracis  
XX wild-type PA protein sequence shown in figure 13 (AAMS1483)  
XX  
XX Sequence 735 AA;  
XX  
XX Query Match 99.0%; Score 3613; DB 5; Length 735;  
XX Best Local Similarity 96.5%; Pred. No. 2e-238;  
XX Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;

QY 1 EVKQENLLNESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENLLNESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSEYTFATSDNHNVTMVDQEVINKASNKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSEYTFATSDNHNVTMVDQEVINKASNKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKKRSTAGTPVDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKKRSTAGTPVDRDN 180  
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240  
DB 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVSFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASTPTTMNNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 456  
DB 421 LNAQKDASTPTTMNNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 480  
QY 457 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDLETTKPDMTLKEALKIAGFNPENGL 516  
DB 481 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDLETTKPDMTLKEALKIAGFNPENGL 540  
QY 517 QYQKDIETDFNPDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 576  
DB 541 QYQKDIETDFNPDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 600  
QY 577 NNTAVGADESVEKAEHREVNSTEGLLNIDKDIRKILSGYVIEDETEGLKEVINDRY 636  
DB 601 NNTAVGADESVEKAEHREVNSTEGLLNIDKDIRKILSGYVIEDETEGLKEVINDRY 660  
QY 637 DMLNISSLRQDKTFFDKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 696  
DB 661 DMLNISSLRQDKTFFDKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 697 IKKILIFSKKGYEIG 711  
DB 721 IKKILIFSKKGYEIG 735

RESULT 10  
AAW51490  
ID AAW51490 standard; protein; 735 AA.  
XX  
AC AAW51490;  
XX  
DT 01-FEB-2002 (first entry)  
DE Anthrax PA mutant D425E.  
XX  
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
OS Bacillus anthracis.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 425 /note= "Wild-type Asp substituted by Glu"

XX WO200182788-A2.  
PN 08-NOV-2001.  
XX  
PD 04-MAY-2001; 2001WO-US014372.  
XX  
PP 04-MAY-2000; 2000US-0201800P.  
XX  
PR (HARD ) HARVARD COLLEGE.  
XX  
PA Collier RJ, Sellman BR;  
XX  
PI WPI; 2002-017725/02.  
XX  
PS Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX  
PS Claim 4; Page; 77pp; English.  
XX  
CC The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAW52113 and AAW51484-AAW51500) especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAW51483)  
XX  
SQ Sequence 735 AA;  
Query Match 98.9%; Score 3609; DB 5; Length 735;  
Best Local Similarity 96.3%; Pred. No. 3.7e-238;  
Matches 708; Conservative 1; Mismatches 2; Indels 24; Gaps 1;  
QY 1 EVKQENLLNESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENLLNESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSEYTFATSDNHNVTMVDQEVINKASNKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSEYTFATSDNHNVTMVDQEVINKASNKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKKRSTAGTPVDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKKRSTAGTPVDRDN 180  
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240  
DB 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVSFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASTPTTMNNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 456  
DB 421 LNAQKDASTPTTMNNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTQSNWSEV 480  
QY 457 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDLETTKPDMTLKEALKIAGFNPENGL 516  
DB 481 LPOIQTETARIIFNGKDLNVERRIAAYVNPSPDLETTKPDMTLKEALKIAGFNPENGL 540  
QY 517 QYQKDIETDFNPDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 576

Db	541	QYQKDIETEDFNDQTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRKRFHYDR	600
Qy	577	NNTAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY	636
Db	601	NNTAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY	660
Qy	637	DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIIINPSENGDTSTNG	696
Db	661	DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIIINPSENGDTSTNG	720
Qy	697	IKKILIFSKKGYEIG 711	
Db	721	IKKILIFSKKGYEIG 735	
RESULT 11			
AA51487	standard; protein; 735 AA.		
XX	AC	AA51487;	
XX	DT	01-FEB-2002 (first entry)	
XX	DE	Anthrax PA mutant K397Q.	
XX	XX	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;	
XX	KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.	
XX	OS	Bacillus anthracis.	
XX	OS	Synthetic.	
XX	Key	Location/Qualifiers	
XX	FT	Misc-difference 397	
XX	FT	/note= "Wild-type Lys substituted by Gln"	
XX	PN	WO200182788-A2.	
XX	XX	08-NOV-2001.	
XX	XX	04-MAY-2001; 2001WO-US014372.	
XX	PR	04-MAY-2000; 2000US-0201800P.	
XX	XX	(HARD ) HARVARD COLLEGE.	
XX	XX	Collier RJ, Sellman BR;	
XX	XX	WPI; 2002-017725/02.	
XX	PT	Protecting humans against anthrax using mutant B groups (anthrax	
XX	PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus	
XX	XX	anthracis.	
XX	PS	Claim 4; Page; 77pp; English.	
XX	XX	The invention relates to antibacterial agents comprising mutant forms of	
XX	CC	pore-forming toxins (AA52113 and AA51484-AA51500), especially mutants	
XX	CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B	
XX	CC	moiety is anthrax protective antigen (PA) and using these mutants or	
XX	CC	compositions of them for protecting against Bacillus anthracis infections	
XX	CC	in humans, especially as vaccines. Note: The present sequence is not	
XX	CC	given in the specification but is derived from the Bacillus anthracis	
XX	CC	wild-type PA protein sequence shown in figure 13 (AA51483)	
XX	SQ	Sequence 735 AA;	
Query Match 98.9%; Score 3608; DB 5; Length 735;			
Best Local Similarity 96.3%; Pred. No. 4.4e-238;			
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;			
Qy	1	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYQF	60

Db	1	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYQF	60
Qy	61	QSAIWSGFIKVKSDDEVTFATSDADNHTVMWDDQDEVINKASNSNKIRLEKGRLYQIKIY	120
Db	61	QSAIWSGFIKVKSDDEVTFATSDADNHTVMWDDQDEVINKASNSNKIRLEKGRLYQIKIY	120
Qy	121	QRENTEKGLDFKLYWTDSONKKEVISSDNIQLPELKQKSSNSKRGKSTAGTPVPDRN	180
Db	121	QRENTEKGLDFKLYWTDSONKKEVISSDNIQLPELKQKSSNSKRGKSTAGTPVPDRN	180
Qy	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEKT	240
Db	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKMWSTASDPYDFEKT	240
Qy	241	GRIDKNVSPARHPLVAAAPVIVHVDMENIISKNEQSTQNTDGETRISKNTSTSRHT	300
Db	241	GRIDKNVSPARHPLVAAAPVIVHVDMENIISKNEQSTQNTDGETRISKNTSTSRHT	300
Qy	301	-----SNSNSTVAIDHLSLAGERTWAETMGLNTADTARL	336
Db	301	SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL	360
Qy	337	NANIRYVNTGTAPIYNNVLTPTSLVLGKQTLATIKADENOLSOILLAPNNYPSKXLAPIA	396
Db	361	NANIRYVNTGTAPIYNNVLTPTSLVLGKQTLATIKADENOLSOILLAPNNYPSKXLAPIA	420
Qy	397	LNAQKDASPTITMNVNPFLEKTKQLRLDQDQYGNIAFYNFENGRVVDVTSNWSSEV	456
Db	421	LNAQDFFSPTITMNVNPFLEKTKQLRLDQDQYGNIAFYNFENGRVVDVTSNWSSEV	480
Qy	457	LPOIQTETARIIFNGKOLNIVERRIAAVNPSPLETTKPDMLKEALKIAGFNPNGNL	516
Db	481	LPOIQTETARIIFNGKOLNIVERRIAAVNPSPLETTKPDMLKEALKIAGFNPNGNL	540
Qy	517	OYQKDIETEDFNDQTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRKRFHYDR	576
Db	541	OYQKDIETEDFNDQTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRKRFHYDR	600
Qy	577	NNI-AVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY	636
Db	601	NNI-AVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDTEGLKEVINDRY	660
Qy	637	DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIIINPSENGDTSTNG	696
Db	661	DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVYAVTKENTIIINPSENGDTSTNG	720
Qy	697	IKKILIFSKKGYEIG 711	
Db	721	IKKILIFSKKGYEIG 735	
RESULT 12			
AA51500	standard; protein; 735 AA.		
XX	AC	AA51500;	
XX	DT	01-FEB-2002 (first entry)	
XX	DE	Anthrax PA F427K mutant.	
XX	KW	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;	
XX	KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.	
XX	OS	Bacillus anthracis.	
XX	OS	Synthetic.	
XX	Key	Location/Qualifiers	
XX	FT	Misc-difference 427	
XX	FT	/note= "Wild-type Phe substituted by Lys"	
XX	PN	WO200182788-A2.	

PD 08-NOV-2001.  
 XX  
 PF 04-MAY-2001; 2001WO-US014372.  
 XX  
 PR 04-MAY-2000; 2000US-0201800P.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Collier RJ, Sellman BR;  
 XX  
 DR WPI; 2002-017725/02.  
 XX  
 XX  
 PT Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of *Bacillus*  
 PT anthracis.  
 XX  
 XX  
 PS Claim 4; Page; 77pp; English.  
 XX  
 XX  
 CC The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against *Bacillus anthracis* infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the *Bacillus anthracis*  
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
 XX  
 XX  
 SQ Sequence 735 AA;  
 Query Match 98.9%; Score 3608; DB 5; Length 735;  
 Best Local Similarity 96.3%; Pred. No. 4.4e-238;  
 Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
 QY 1 EVKQENRLNESSSQGLLGYYFSDLNFCAPMVVTSSTGDLSPSSSELENIPSENOYF 60  
 DB 1 EVKQENRLNESSSQGLLGYYFSDLNFCAPMVVTSSTGDLSPSSSELENIPSENOYF 60  
 QY 61 QSAIWSGFIKVKSDYEYFATSADNHTVMVDQEVINKASNSKIRLEKGLYQIKIY 120  
 DB 61 QSAIWSGFIKVKSDYEYFATSADNHTVMVDQEVINKASNSKIRLEKGLYQIKIY 120  
 QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQSSNSRKKRSTASGPTVPDRDN 180  
 DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQSSNSRKKRSTASGPTVPDRDN 180  
 QY 181 DGIPDSLEVEGYTVVKNKFTPLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240  
 DB 181 DGIPDSLEVEGYTVVKNKFTPLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240  
 QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
 DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
 QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
 DB 301 SEVHGNAEVHASFDPDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 QY 337 NANIRYNTGTAPYVNLPTTSLVLGNQOTLATIKADENQLSQILAPNNYPSKNLAPTA 396  
 DB 361 NANIRYNTGTAPYVNLPTTSLVLGNQOTLATIKADENQLSQILAPNNYPSKNLAPTA 420  
 QY 397 LNAQKQASSTPTWNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 456  
 DB 421 LNAQKQASSTPTWNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 480  
 QY 457 LPQIQETARTIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLKEALKAFGNPNGL 516  
 DB 481 LPQIQETARTIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLKEALKAFGNPNGL 540  
 QY 517 QYQKQITEFDNFQDQTSQNKQLAELNATNIYVLDKIKUNAKQNTILIRDKRPHYDR 576  
 DB 541 QYQKQITEFDNFQDQTSQNKQLAELNATNIYVLDKIKUNAKQNTILIRDKRPHYDR 600

QY 577 NNIAVGADESVVKEAHREVINSSTEGLLNIDKDIRKLSGYIVEIEDTGLKKEVINDRY 636  
 DB 601 NNIAVGADESVVKEAHREVINSSTEGLLNIDKDIRKLSGYIVEIEDTGLKKEVINDRY 660  
 QY 637 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVNVAVTKEHTIINPSENGDTSTNG 696  
 DB 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVNVAVTKEHTIINPSENGDTSTNG 720  
 QY 697 IKKILIFSKKGYEIG 711  
 DB 721 IKKILIFSKKGYEIG 735  
 RESULT 13  
 AAM51489  
 ID AAM51489 standard; protein; 735 AA.  
 XX  
 XX AAM51489;  
 XX 01-FEB-2002 (first entry)  
 XX Anthrax PA mutant D425N.  
 XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 XX B moiety; A-B anthrax toxin; *Bacillus anthracis*; vaccine; mutant; muten.  
 XX *Bacillus anthracis*.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 425  
 FT /note= "Wild-type Asp substituted by Asn"  
 XX  
 XX WC200182788-A2.  
 XX 08-NOV-2001.  
 XX 04-MAY-2001; 2001WO-US014372.  
 XX 04-MAY-2000; 2000US-0201800P.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Collier RJ, Sellman BR;  
 XX WPI; 2002-017725/02.  
 XX  
 XX Protecting humans against anthrax using mutant B groups (anthrax  
 XX protective antigens) of the pore-forming binary A-B toxin of *Bacillus*  
 XX anthracis.  
 XX  
 XX Claim 4; Page; 77pp; English.  
 XX  
 XX The invention relates to antibacterial agents comprising mutant forms of  
 XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 XX moiety is anthrax protective antigen (PA) and using these mutants or  
 XX compositions of them for protecting against *Bacillus anthracis* infections  
 XX in humans, especially as vaccines. Note: The present sequence is not  
 XX given in the specification but is derived from the *Bacillus anthracis*  
 XX wild-type PA protein sequence shown in figure 13 (AAM51483)  
 XX  
 XX  
 SQ Sequence 735 AA;  
 Query Match 98.9%; Score 3608; DB 5; Length 735;  
 Best Local Similarity 96.3%; Pred. No. 4.4e-238;  
 Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
 QY 1 EVKQENRLNESSSQGLLGYYFSDLNFCAPMVVTSSTGDLSPSSSELENIPSENOYF 60  
 DB 1 EVKQENRLNESSSQGLLGYYFSDLNFCAPMVVTSSTGDLSPSSSELENIPSENOYF 60  
 QY 61 QSAIWSGFIKVKSDYEYFATSADNHTVMVDQEVINKASNSKIRLEKGLYQIKIY 120  
 DB 61 QSAIWSGFIKVKSDYEYFATSADNHTVMVDQEVINKASNSKIRLEKGLYQIKIY 120  
 QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQSSNSRKKRSTASGPTVPDRDN 180  
 DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQSSNSRKKRSTASGPTVPDRDN 180  
 QY 181 DGIPDSLEVEGYTVVKNKFTPLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240  
 DB 181 DGIPDSLEVEGYTVVKNKFTPLSPWISNIHEKKGLTKYKSSPEKMWSTASDPYDFEKT 240  
 QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
 DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
 QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
 DB 301 SEVHGNAEVHASFDPDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 QY 337 NANIRYNTGTAPYVNLPTTSLVLGNQOTLATIKADENQLSQILAPNNYPSKNLAPTA 396  
 DB 361 NANIRYNTGTAPYVNLPTTSLVLGNQOTLATIKADENQLSQILAPNNYPSKNLAPTA 420  
 QY 397 LNAQKQASSTPTWNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 456  
 DB 421 LNAQKQASSTPTWNNYQFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 480  
 QY 457 LPQIQETARTIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLKEALKAFGNPNGL 516  
 DB 481 LPQIQETARTIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLKEALKAFGNPNGL 540  
 QY 517 QYQKQITEFDNFQDQTSQNKQLAELNATNIYVLDKIKUNAKQNTILIRDKRPHYDR 576  
 DB 541 QYQKQITEFDNFQDQTSQNKQLAELNATNIYVLDKIKUNAKQNTILIRDKRPHYDR 600

Db 61 QSAIWSGFIKVKSDYFTFATSADNHYTMWVDDQVINKASNSKIRLEKGRLYQIKQY 120  
 QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQKSSNRKSRSTAGTVPDRDN 180  
 Db 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQKSSNRKSRSTAGTVPDRDN 180  
 QY 181 DGIPLDSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASTADPYDFEKT 240  
 Db 181 DGIPLDSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASTADPYDFEKT 240  
 QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
 Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
 QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
 Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
 Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
 QY 397 LNAQKASSTPTITMNNQFLEBKTQKRLDITDQVYGNATYNFENGVRVDTGNSWSEV 456  
 Db 421 LNAQKASSTPTITMNNQFLEBKTQKRLDITDQVYGNATYNFENGVRVDTGNSWSEV 480  
 QY 457 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETKPDMTLKEALKIARFNEPKNL 516  
 Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETKPDMTLKEALKIARFNEPKNL 540  
 QY 517 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLAKMILLIRDKRFHYDR 576  
 Db 541 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLAKMILLIRDKRFHYDR 600  
 QY 577 NNIAVGADESUVKAEHREVINSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 636  
 Db 601 NNIAVGADESUVKAEHREVINSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 660  
 QY 637 DMLNISLRQDGTFFDKKYNKLPYISNPNYKNNYAVTKENTIINPSENGDTSTNG 696  
 Db 661 DMLNISLRQDGTFFDKKYNKLPYISNPNYKNNYAVTKENTIINPSENGDTSTNG 720  
 QY 697 IKKILIFSKKGYEIG 711  
 Db 721 IKKILIFSKKGYEIG 735

## RESULT 14

AAR60179  
 ID AAR60179 standard; protein; 735 AA.

AC AAR60179;  
 XX

XX 25-MAR-2003 (revised)  
 DT 03-APR-1995 (first entry)

XX  
 XX

Protective antigen of *Bacillus anthracis*.

XX Anthrax; *Bacillus anthracis*; fusion protein; protective antigen;  
 KW protective antigen; cell killing; targeting; pathogen;  
 KW intracellular; HIV; human immunodeficiency virus; toxin.

XX *Bacillus anthracis*.

XX WO9418332-A2.

XX 18-AUG-1994.

XX 14-FEB-1994; 94WO-US001624.

XX 12-FEB-1993; 93US-00021601.

XX 25-JUN-1993; 93US-00082849.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX Leppla SH, Klimpel K, Arora N, Singh Y, Nichols PJ;  
 XX WFI; 1994-279753/34.  
 DR N-PSDB; AAQ70180.  
 XX  
 PT Nucleic acid encoding anthrax toxin fusion protein - useful for  
 PT targeting toxin to specific cells, eg for killing tumour cells or HIV-  
 PT infected cells.  
 XX  
 PS Disclosure; Page 81-83; 124pp; English.  
 XX  
 CC The sequence encoding the protective antigen of *Bacillus anthracis* may be  
 CC used in the construction of a nucleic acid which encodes a fusion protein  
 CC comprising the anthrax protective antigen binding domain of the native  
 CC anthrax lethal factor and a sequence encoding an activity inducing domain  
 CC of a second protein. The fusion proteins are useful for the specific  
 CC killing of tumour cells or the killing of cells infected with  
 CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX  
 SQ Sequence 735 AA;

Query Match 98.9%; Score 3607; DB 2; Length 735;  
 Best Local Similarity 96.3%; Pred. No. 5,1e-238;  
 Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNESESSQGLLYFSDLNFPQPMVVTSTTGLDIPSELENISENQYF 60  
 Db 1 EVKQENRLNESESSQGLLYFSDLNFPQPMVVTSTTGLDIPSELENISENQYF 60  
 QY 61 QSAIWSGFIKVKSDYFTFATSADNHYTMWVDDQVINKASNSKIRLEKGRLYQIKQY 120  
 Db 61 QSAIWSGFIKVKSDYFTFATSADNHYTMWVDDQVINKASNSKIRLEKGRLYQIKQY 120  
 QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQKSSNRKSRSTAGTVPDRDN 180  
 Db 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQKSSNRKSRSTAGTVPDRDN 180  
 QY 181 DGIPLDSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASTADPYDFEKT 240  
 Db 181 DGIPLDSLEVGTYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWSASTADPYDFEKT 240  
 QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
 Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
 QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
 Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 QY 337 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
 Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
 QY 397 LNAQKASSTPTITMNNQFLEBKTQKRLDITDQVYGNATYNFENGVRVDTGNSWSEV 456  
 Db 421 LNAQKASSTPTITMNNQFLEBKTQKRLDITDQVYGNATYNFENGVRVDTGNSWSEV 480  
 QY 457 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETKPDMTLKEALKIARFNEPKNL 516  
 Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETKPDMTLKEALKIARFNEPKNL 540  
 QY 517 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLAKMILLIRDKRFHYDR 576  
 Db 541 QYQKDIETFDNFDOQTSONIKQKLAELNATNIYVLDKIKLAKMILLIRDKRFHYDR 600  
 QY 577 NNIAVGADESUVKAEHREVINSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 636  
 Db 601 NNIAVGADESUVKAEHREVINSTEGLLNIDKDIRKILSGYIVIEDETEGLKEVINDRY 660

QY 637 DMLNLSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 696  
DB 661 DMLNLSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 697 IKKILIFSKGYEIG 711  
DB 721 IKKILIFSKGYEIG 735

RESULT 15  
ID AAM51483  
XX AAM51483 standard; protein; 735 AA.  
XX  
AC AAM51483;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Anthrax PA protein.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.  
XX  
OS Bacillus anthracis.  
XX  
PN WO200182788-A2.  
XX  
XX 08-NOV-2001.  
XX  
PF 04-MAY-2001; 2001WO-US014372.  
XX  
PR 04-MAY-2000; 2000US-0201800P.  
XX  
XX (HARD) HARVARD COLLEGE.  
XX  
XX Collier RJ, Sellman BR;  
PI  
XX WPI; 2002-017725/02.  
DR N-PSDB; AA199904.  
XX  
PT Protecting humans against anthrax using mutant B groups (anthrax  
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
PT anthracis.  
XX  
PS Disclosure; Fig 13; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. The present sequence is that of the  
CC anthrax PA protein  
XX  
SQ Sequence 735 AA;

Query Match 98.9%; Score 3607; DB 5; Length 735;  
Best Local Similarity 96.3%; Pred. No. 5.1e-238;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLLYGYSFSLNFOAPMVVTSSTGDLSPSSLENIIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLLYGYSFSLNFOAPMVVTSSTGDLSPSSLENIIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNVHVMVDQEVINKASNKIRLEKGRLYQIKIQY 120  
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNVHVMVDQEVINKASNKIRLEKGRLYQIKIQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRRSTAGTVPDRN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNRKRRSTAGTVPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEYKT 240

DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEYKT 240  
QY 241 GRIDKNVSPFARHPLVAAAPVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
DB 241 GRIDKNVSPFARHPLVAAAPVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERWAEFTMGINTADTARL 336  
DB 301 SEVHGNAEVHASFDFDIGGSVAGFNSNSTVAIDHSLSLAGERWAEFTMGINTADTARL 360  
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILLAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILLAPNNYPSKNLAPIA 420  
QY 397 LNAQKDAASSTPITMNYNOFLELEKTKQLRLDQVYGNIAVYVNFENGVRVDTGNNWSEV 456  
DB 421 LNAQDDFSSTPITMNYNOFLELEKTKQLRLDQVYGNIAVYVNFENGVRVDTGNNWSEV 480  
QY 457 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 516  
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540  
QY 517 QYQKDIETEFDFNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAKMILIRDKPFHYDR 576  
DB 541 QYQKDIETEFDFNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAKMILIRDKPFHYDR 600  
QY 577 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYVIEIETEGKEVINDRY 636  
DB 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYVIEIETEGKEVINDRY 660  
QY 637 DMLNLSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 696  
DB 661 DMLNLSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 697 IKKILIFSKGYEIG 711  
DB 721 IKKILIFSKGYEIG 735

Search completed: May 3, 2004, 19:36:05  
Job time : 49.5168 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 14.6056 Seconds  
(without alignments)  
2513.152 Million cell updates/sec

Title: US-09-848-909A-16  
Perfect score: 3648  
Sequence: 1 EVKQENRLNESESSQGLL.....TSTNGIKLILFSKKGYEIG 711

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

```
Database : Issued Patents AA:*
1: /cgm2_6/pctodara/2/iaa/5A_COMB pep.*
2: /cgm2_6/pctodara/2/iaa/5B_COMB pep.*
3: /cgm2_6/pctodara/2/iaa/6A_COMB pep.*
4: /cgm2_6/pctodara/2/iaa/6B_COMB pep.*
5: /cgm2_6/pctodara/2/iaa/FCTUS COMB pep.*
6: /cgm2_6/pctodara/2/iaa/backfiles1.pe
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query #	Length	DB	ID	Description
1	3607	98.9	735	1	US-08-021-601-4	Sequence 4, Appl
2	3607	98.9	735	1	US-08-082-849B-4	Sequence 4, Appl
3	3607	98.9	735	5	PCU-US94-01624-4	Sequence 4, Appl
4	352.5	97.7	903	1	US-08-021-601-12	Sequence 12, Appl
5	352.5	97.7	903	1	US-08-082-849B-12	Sequence 12, Appl
6	352.5	97.7	903	5	PCU-US94-01624-12	Sequence 12, Appl
7	3430	94.0	719	1	US-08-083-849B-31	Sequence 31, Appl
8	3430	94.0	719	5	PCU-US94-01624-31	Sequence 31, Appl
9	1275	35.0	288	4	US-09-273-839A-8	Sequence 8, Appl
10	759	20.8	881	3	US-08-960-780-32	Sequence 32, Appl
11	759	20.8	881	3	US-09-073-898-32	Sequence 32, Appl
12	759	20.8	881	4	US-09-307-106-8	Sequence 8, Appl
13	759	20.8	881	4	US-09-850-351A-32	Sequence 32, Appl
14	757	20.8	884	1	US-08-471-033-5	Sequence 5, Appl
15	757	20.8	884	2	US-08-471-044-5	Sequence 5, Appl
16	757	20.8	884	2	US-08-463-483A-5	Sequence 5, Appl
17	757	20.8	884	2	US-08-471-046A-5	Sequence 5, Appl
18	757	20.8	884	2	US-08-470-566B-5	Sequence 5, Appl
19	757	20.8	884	2	US-08-469-334-5	Sequence 5, Appl
20	757	20.8	884	3	US-09-300-529-5	Sequence 5, Appl
21	757	20.8	1346	1	US-08-471-033-23	Sequence 23, Appl
22	757	20.8	1346	1	US-08-471-044-23	Sequence 23, Appl
23	757	20.8	1346	2	US-08-463-483A-23	Sequence 23, Appl
24	757	20.8	1346	2	US-08-471-046A-23	Sequence 23, Appl
25	757	20.8	1346	2	US-08-470-566B-23	Sequence 23, Appl
26	757	20.8	1346	2	US-08-469-334-23	Sequence 23, Appl
27	757	20.8	1346	3	US-09-300-529-23	Sequence 23, Appl

## ALIGNMENTS

RESULT 1  
 US-08-021-601-4  
 Sequence 4, Application US/08021601  
 Patent No. 5591631  
 GENERAL INFORMATION:  
 APPLICANT: Leppia, Stephen H.  
 APPLICANT: Klimpel, Kurt R.  
 APPLICANT: Nichols, Peter J.  
 APPLICANT: Azora, Naveen  
 APPLICANT: Singh, Yogendra  
 TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
 TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
 NUMBER OF SEQUENCES: 12  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Needle & Rosenberg, P.C.  
 STREET: 133 Carnegie Way, Suite 400  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: USA  
 ZIP: 30303  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/021,601  
 FILING DATE: 19930212  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Spratt, Gwendolyn D.  
 REGISTRATION NUMBER: 36,016  
 REFERENCE/DOCKET NUMBER: 1414.057  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 404/588-0770  
 TELEFAX: 404/588-9880  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 735 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-021-601-4

Query Match: 98.9%; Score 3607; DB 1; Length 735;  
Best Local Similarity 96.3%; Pred. No. 6.2e-261;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
1 EVKQENRLNSESSSGCLGYYFSDLNFPQMVVTSSTGDLSPSELNIPSENQYF 60

1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSRKRSTAGPTVPDRN 180  
121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSRKRSTAGPTVPDRN 180  
181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFPKVT 240  
181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFPKVT 240  
241 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
241 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
337 NANIRYVNTGTAPIYVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
361 NANIRYVNTGTAPIYVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
397 LNAQKDAASSTPITMNYNQFLELEKTQKRLDQVYGNIAATYFNGRVRVDTGNSWSEV 456  
421 LNAQKDAASSTPITMNYNQFLELEKTQKRLDQVYGNIAATYFNGRVRVDTGNSWSEV 480  
457 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 516  
481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
517 QYQKDIITEFDNFDOQTQSONKNQLAELNATNIYVLDKIKNAKNMILIRDKRFHYDR 576  
541 QYQKDIITEFDNFDOQTQSONKNQLAELNATNIYVLDKIKNAKNMILIRDKRFHYDR 600  
577 NNIAVGADES VVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 636  
601 NNIAVGADES VVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

RESULT 2  
US-08-082-849B-4  
; Sequence 4, Application US/08082849B  
; Patent No. 5677274  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
; TITLE OF INVENTION: Related Methods  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082.849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-4  
Query Match 98.9%; Score 3607; DB 1; Length 735;  
Best Local Similarity 96.3%; Pred. No. 6.2e-261;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDSEYTFATSAHNVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSRKRSTAGPTVPDRN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSNRSRKRSTAGPTVPDRN 180  
QY 181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFPKVT 240  
DB 181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFPKVT 240  
QY 241 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
DB 241 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDAASSTPITMNYNQFLELEKTQKRLDQVYGNIAATYFNGRVRVDTGNSWSEV 456  
DB 421 LNAQKDAASSTPITMNYNQFLELEKTQKRLDQVYGNIAATYFNGRVRVDTGNSWSEV 480  
QY 457 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 516  
DB 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
QY 517 QYQKDIITEFDNFDOQTQSONKNQLAELNATNIYVLDKIKNAKNMILIRDKRFHYDR 576  
DB 541 QYQKDIITEFDNFDOQTQSONKNQLAELNATNIYVLDKIKNAKNMILIRDKRFHYDR 600  
QY 577 NNIAVGADES VVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 636  
DB 601 NNIAVGADES VVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660



QY 637 DMLNLSLRDQKTFIDFKYNDKPLIYISPNYKVVAVTKENTIIINPSENGDTSTNG 696  
DB 661 DMLNLSLRDQKTFIDFKYNDKPLIYISPNYKVVAVTKENTIIINPSENGDTSTNG 720

QY 697 IKKILIPSKGYEIG 711  
DB 721 IKKILIPSKGYEIG 735

RESULT 3  
PCT-US94-01624-4  
; Sequence 4, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppa, Stephen H.  
; APPLICANT: Klmpel, Kurt R.  
; APPLICANT: Arota, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW  
; STREET: Stewart Street Tower, 20th Floor, One Market  
; STREET: Plaza  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01624  
; FILING DATE: June 25, 1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US94-01624-4

Query Match 98.9%; Score 3607; DB 5; Length 735;  
Best Local Similarity 96.3%; Pred. No. 6.2e-261;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNESSSQGLGYFFDLNFQAPMVVTSSTGDLSPSSSELENIPSENYF 60  
DB 1 EVKQENRLNESSSQGLGYFFDLNFQAPMVVTSSTGDLSPSSSELENIPSENYF 60

QY 61 QSAIWSGFIVKKSDEYTFATSDADNHVWDDQEVINKASNSKIRLEKGRLYQIKQY 120  
DB 61 QSAIWSGFIVKKSDEYTFATSDADNHVWDDQEVINKASNSKIRLEKGRLYQIKQY 120

QY 121 QRENPTKGLDFKLWYTDSONKKEVTSDDNLQLPKQKSSNSRKRSTTSAGTVPDRDN 180  
DB 121 QRENPTKGLDFKLWYTDSONKKEVTSDDNLQLPKQKSSNSRKRSTTSAGTVPDRDN 180

QY 181 DGIIPDSLEVGTYVDVKNKRTFLSPWISNTHKGLTKYKSSPEKSTADDPYSDFEKVT 240  
DB 181 DGIIPDSLEVGTYVDVKNKRTFLSPWISNTHKGLTKYKSSPEKSTADDPYSDFEKVT 240

QY 241 GRIDKNVSPEARHPLVAAPYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPEARHPLVAAPYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTSRHT 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 336  
DB 301 SEVHCNAEVSASFDPDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360

QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQSLQILAPNNYPPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQSLQILAPNNYPPSKNLAPIA 420

QY 397 LNAQKDSASTPTTNNYNQFLELEKTKQLRLDQVYGNIAFYNFENGVRVDTGSNWSEV 456  
DB 421 LNAQKDSASTPTTNNYNQFLELEKTKQLRLDQVYGNIAFYNFENGVRVDTGSNWSEV 480

QY 457 LPOIETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 516  
DB 481 LPOIETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540

QY 517 QYQKDIITEFDNFQOQSQNIKNQLAELNATNIYVLDKIKLNAKMNILIRDKRFHYDR 576  
DB 541 QYQKDIITEFDNFQOQSQNIKNQLAELNATNIYVLDKIKLNAKMNILIRDKRFHYDR 600

QY 577 NNIAVGADSVVKEAHREVINSSTGLLLNDKDKIRKILSGYIVIEDETEGLKEVINDRY 636  
DB 601 NNIAVGADSVVKEAHREVINSSTGLLLNDKDKIRKILSGYIVIEDETEGLKEVINDRY 660

QY 637 DMLNLSLRDQKTFIDFKYNDKPLIYISPNYKVVAVTKENTIIINPSENGDTSTNG 696  
DB 661 DMLNLSLRDQKTFIDFKYNDKPLIYISPNYKVVAVTKENTIIINPSENGDTSTNG 720

QY 697 IKKILIPSKGYEIG 711  
DB 721 IKKILIPSKGYEIG 735

RESULT 4  
US-08-021-601-12  
; Sequence 12, Application US/08021601  
; Patent No. 5591631  
; GENERAL INFORMATION:  
; APPLICANT: Leppa, Stephen H.  
; APPLICANT: Klmpel, Kurt R.  
; APPLICANT: Nichols, Peter J.  
; APPLICANT: Arota, Naveen  
; APPLICANT: Singh, Yogendra  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Needle & Rosenberg, P.C.  
; STREET: 133 Carnegie Way, Suite 400  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,601  
; FILING DATE: 19930212  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 1414,057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-021-601-12

Query Match 97.7%; Score 3562.5; DB 1; Length 903;  
Best Local Similarity 95.5%; Pred. No. 1.8e-257;  
Matches 701; Conservative 1; Mismatches 5; Indels 27; Gaps 2;  
Qy 1 EVKQENLLNESSESSQGLLYGFFDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENLLNESSESSQGLLYGFFDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKSDXYTFAATADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDXYTFAATADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKSTASGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKSTASGPTVPDRN 180  
Qy 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEY 240  
Db 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEY 240  
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVEHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Qy 397 LNAQKQDASPTITMNVNQFLEKTKQLRLDTPQVYGNATYFNFENGVRVDTGNSWSEV 456  
Db 421 LNAQKQDASPTITMNVNQFLEKTKQLRLDTPQVYGNATYFNFENGVRVDTGNSWSEV 480  
Qy 457 LQIQSTTARIIFNGKDLNLVERRIAAANVPSDPLETTKPDMTLKEALKIAGFNEPQNL 516  
Db 481 LQIQSTTARIIFNGKDLNLVERRIAAANVPSDPLETTKPDMTLKEALKIAGFNEPQNL 540  
Qy 517 QYQKDIITFDNFDOOTSONIKNQLAELNATNIYVLDKIKNAKMLIRDKRFFHYDR 576  
Db 541 QYQKDIITFDNFDOOTSONIKNQLAELNATNIYVLDKIKNAKMLIRDKRFFHYDR 600  
Qy 577 NNIAGADESVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 636  
Db 601 NNIAGADESVKAEHREVINSTEGLLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
Qy 637 DMLNSSLQDQGTFTDFKKYNDKPLYSINPNYKVVAVYKNTIINPSNGTSTNG 696  
Db 661 DMLNSSLQDQGTFTDFKKYNDKPLYSINPNYKVVAVYKNTIINPSNGTSTNG 720  
Qy 697 IKKIL---IFSKKG 707  
Db 721 IKKILKKVGLGKG 734

RESULT 5  
US-08-082-849B-12  
Sequence 12, Application US/08082849B  
Patent No. 5677274  
GENERAL INFORMATION:  
APPLICANT: Leppa, Stephen H.  
APPLICANT: Klumpel, Kurt R.  
APPLICANT: Aurora, Naveen

APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
TITLE OF INVENTION: Related Methods  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-12

Query Match 97.7%; Score 3562.5; DB 1; Length 903;  
Best Local Similarity 95.5%; Pred. No. 1.8e-257;  
Matches 701; Conservative 1; Mismatches 5; Indels 27; Gaps 2;  
Qy 1 EVKQENLLNESSESSQGLLYGFFDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENLLNESSESSQGLLYGFFDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKSDXYTFAATADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDXYTFAATADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKSTASGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKSTASGPTVPDRN 180  
Qy 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEY 240  
Db 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEY 240  
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVEHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Qy 397 LNAQKQDASPTITMNVNQFLEKTKQLRLDTPQVYGNATYFNFENGVRVDTGNSWSEV 456

Db 421 LNAQDFSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSGNWSEV 480  
QY 457 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 516  
Db 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 540  
QY 517 QYQKDIETEFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKVNILIRKRPHYDR 576  
Db 541 QYQKDIETEFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKVNILIRKRPHYDR 600  
QY 577 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 636  
Db 601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
QY 637 DMLNISSLRQDGKTFIDFKKYNLPLYSNPYKVNIAATYFNGRVRVDTGSGNWSEV 696  
Db 661 DMLNISSLRQDGKTFIDFKKYNLPLYSNPYKVNIAATYFNGRVRVDTGSGNWSEV 720  
QY 697 IKKIL---IFSCKG 707  
Db 721 IKKILKVVGLGKG 734  
RESULT 6  
PCT-US94-01624-12  
; Sequence 12, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW  
; STREET: Steuart Street Tower, 20th Floor, One Market  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01624  
; FILING DATE: June 25, 1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-115  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 903 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-01624-12  
Query Match 97.7%; Score 3562.5; DB 5; Length 903;  
Best Local Similarity 95.5%; Pred. No. 1.8e-257;  
Matches 701; Conservative 1; Mismatches 5; Indels 27; Gaps 2;  
QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQVF 60

Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQVF 60  
QY 61 QSAIWSGFIKVKSDSEYTFATSDNHVMTWDDQEVINKASNSKIRLEKGLYQIKQY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSDNHVMTWDDQEVINKASNSKIRLEKGLYQIKQY 120  
QY 121 QRENTEKGLDFKLWTDSONKKEVSSDNLQQLPELKQKSSNSRKRKSTSGPTVPDRDN 180  
Db 121 QRENTEKGLDFKLWTDSONKKEVSSDNLQQLPELKQKSSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240  
QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENOLSOILAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKADENOLSOILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSGNWSEV 456  
Db 421 LNAQKDASTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSGNWSEV 480  
QY 457 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 516  
Db 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDWTLKEALKIAFGNEPENG 540  
QY 517 QYQKDIETEFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKVNILIRKRPHYDR 576  
Db 541 QYQKDIETEFDFNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKVNILIRKRPHYDR 600  
QY 577 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 636  
Db 601 NNIAGDAESVVKAEHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
QY 637 DMLNISSLRQDGKTFIDFKKYNLPLYSNPYKVNIAATYFNGRVRVDTGSGNWSEV 696  
Db 661 DMLNISSLRQDGKTFIDFKKYNLPLYSNPYKVNIAATYFNGRVRVDTGSGNWSEV 720  
QY 697 IKKIL---IFSCKG 707  
Db 721 IKKILKVVGLGKG 734  
RESULT 7  
US-08-082-849B-31  
; Sequence 31, Application US/08082849B  
; Patent No. 5677274  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
; TITLE OF INVENTION: Related Methods  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082.849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021.601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-31

Query Match 94.0%; Score 3430; DB 1; Length 719;  
Best Local Similarity 91.9%; Pred. No. 9.8e-248;  
Matches 679; Conservative 4; Mismatches 8; Indels 48; Gaps 3;

QY 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
DB 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
QY 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS-----RKKSTSGPTVP 176  
DB 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSNTATIMQKGNFLOGPTVP 180  
QY 177 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYSD 236  
DB 181 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYSD 240  
QY 237 EKVTRIDKNSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTS 296  
DB 241 EKVTRIDKNSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTS 300  
QY 297 RTHT-----SNSNSTVAIDHSLSLAGERWAEWTGMLNTAD 332  
DB 301 RTHTSEVHGNAEVAHFFDIGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTGMLNTAD 360  
QY 333 TARNANIRVNTGTAPIYVNLPTTSLVJGKQTLATIKADENQSLQILAPNYPKSL 392  
DB 361 TARNANIRVNTGTAPIYVNLPTTSLVJGKQTLATIKADENQSLQILAPNYPKSL 420  
QY 393 APIALNAQDKASPTITMNYNQFLEKTKQLRLDDQVYGNATYVFNENGRVVDTSN 452  
DB 421 APIALNAQDDFSSTPTMNY-----YGNATYVFNENGRVVDTSN 460  
QY 453 WSEVLPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKALKIATGFNEP 512  
DB 461 WSEVLPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKALKIATGFNEP 520  
QY 513 NGNLOYGKDOIETFDNFDOQTQNTKNQALAEINATNYIVLTDKILNANMILLIRDKRF 572  
DB 521 NGNLOYGKDOIETFDNFDOQTQNTKNQALAEINATNYIVLTDKILNANMILLIRDKRF 580  
QY 573 HYDRNNIAGADESVVKEAHREVINSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVI 632  
DB 581 HYDRNNIAGADESVVKEAHREVINSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVI 640

QY 633 NDRYDMLNISSLRQDKTFIDFKYNDKLPYISNPYKVVYAVTKNTIINPSENGDT 692  
DB 641 NDRYDMLNISSLRQDKTFIDFKYNDKLPYISNPYKVVYAVTKNTIINPSENGDT 700  
QY 693 STNGIKKILIFSKKGYEIG 711  
DB 701 STNGIKKILIFSKKGYEIG 719  
RESULT 8  
PCT-US94-01624-31  
Sequence 31, Application PC/TUS9401624  
GENERAL INFORMATION:  
APPLICANT: Leppia, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
TITLE OF INVENTION: RELATED METHODS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOMSEND and TOMSEND KHOURIE and CREW  
STREET: Stewart Street Tower, 20th Floor, One Market  
STREET: Plaza  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/01624  
FILING DATE: June 25, 1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION/DOCKET NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01624-31

Query Match 94.0%; Score 3430; DB 5; Length 719;  
Best Local Similarity 91.9%; Pred. No. 9.8e-248;  
Matches 679; Conservative 4; Mismatches 8; Indels 48; Gaps 3;

QY 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
DB 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
QY 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS-----RKKSTSGPTVP 176  
DB 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSNTATIMQKGNFLOGPTVP 180  
QY 177 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYSD 236  
DB 181 DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYSD 240

QY 237 EKVTRIDKVNUPBARPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNSTIS 296  
 Db 241 EKVTRIDKVNUPBARPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNSTIS 300  
 QY 297 RTHT-----SNSNSTVAIDHSLSLAGERTWAETWGLNTAD 332  
 Db 301 RTHTSEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTAD 360  
 QY 333 TARLNANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNL 392  
 Db 361 TARLNANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNL 420  
 QY 393 APIALNAQKASSTPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNL 452  
 Db 421 APIALNAQKASSTPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNL 460  
 QY 453 WSEVLPQIETARTIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGENEP 512  
 Db 461 WSEVLPQIETARTIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGENEP 520  
 QY 513 NGNLOYOGKDIETEDFNFDOQTSONIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRF 572  
 Db 521 NGNLOYOGKDIETEDFNFDOQTSONIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRF 580  
 QY 573 HYDRNNAVGADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVI 632  
 Db 581 HYDRNNAVGADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVI 640  
 QY 633 NDRYDMLNMISSRQDKTFIDPKYNDKPLIYISNPNYKVNVAVTENTIIINPSENGDT 692  
 Db 641 NDRYDMLNMISSRQDKTFIDPKYNDKPLIYISNPNYKVNVAVTENTIIINPSENGDT 700  
 QY 693 STNGIKKILIFSKGYEIG 711  
 Db 701 STNGIKKILIFSKGYEIG 719

RESULT 9  
 US-09-273-839A-8  
 ; Sequence 8, Application US/09273839A  
 ; Patent No. 6329156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cirino, Nick M  
 ; APPLICANT: Jackson, Paul J  
 ; APPLICANT: Lehnert, Bruce E  
 ; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface  
 ; FILE REFERENCE: S-89,662  
 ; CURRENT APPLICATION NUMBER: US/09/273,839A  
 ; CURRENT FILING DATE: 1999-03-22  
 ; NUMBER OF SEQ ID NOS: 12  
 ; SOFTWARE: Patent in Ver. 2.0  
 ; SEQ ID NO 8  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus anthracis  
 US-09-273-839A-8

Query Match 35.0%; Score 1275; DB 4; Length 288;  
 Best Local Similarity 100.0%; Pred. No. 2e-87;  
 Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 462 ETARTIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGENEPNGNLOYOGK 521  
 Db 29 ETARTIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKAFGENEPNGNLOYOGK 88  
 QY 522 DITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRFHYDRNNAV 581  
 Db 89 DITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRFHYDRNNAV 148  
 QY 582 GADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNI 641  
 Db 149 GADESIVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRYDMLNI 208

QY 642 SSLROGDKTFIDPKYNDKPLIYISNPNYKVNVAVTENTIIINPSENGDTSTNGIKKIL 701  
 Db 209 SSLROGDKTFIDPKYNDKPLIYISNPNYKVNVAVTENTIIINPSENGDTSTNGIKKIL 268  
 QY 702 IFSKGYEIG 711  
 Db 269 IFSKGYEIG 278  
 RESULT 10  
 US-08-960-780-32  
 ; Sequence 32, Application US/08960780  
 ; Patent No. 6204435  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Feitelson, Gerald S.  
 ; APPLICANT: Schnepf, H. Ernest  
 ; APPLICANT: Narve, Kenneth E.  
 ; APPLICANT: Stockhoff, Brian A.  
 ; APPLICANT: Schmeits, James  
 ; APPLICANT: Loewer, David  
 ; APPLICANT: Dullum, Charles Joseph  
 ; APPLICANT: Muller-Cohn, Judy  
 ; APPLICANT: Stamp, Lisa  
 ; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide  
 ; TITLE OF INVENTION: Sequences Which Encode These Toxins  
 ; NUMBER OF SEQUENCES: 134  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: FL  
 ; COUNTRY: US  
 ; ZIP: 32606-6669  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/960,780  
 ; FILING DATE: 30-OCT-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/029,848  
 ; FILING DATE: 30-OCT-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Saliwanchik, David R.  
 ; REGISTRATION NUMBER: 31,794  
 ; REFERENCE/DOCKET NUMBER: MA-708  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 352-375-8100  
 ; TELEFAX: 352-372-5800  
 ; INFORMATION FOR SEQ ID NO: 32:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 881 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: Peptide  
 ; ORIGINAL SOURCE:  
 ; INDIVIDUAL ISOLATE: 177C8  
 ; US-08-960-780-32

Query Match 20.8%; Score 759; DB 3; Length 881;  
 Best Local Similarity 29.7%; Pred. No. 3.5e-48;  
 Matches 236; Conservative 136; Mismatches 294; Indels 128; Gaps 32;  
 QY 4 QENRLNSESQGLGYFFSDLNIFQAPNVVSTTGDLSIPSSLEN--IPSENYQYQ 61  
 Db 39 QXNQ---QKMDRKLGLGYFFKQKDF-SNLTMEFAPTRDSTLIYDQQTANKLLDKQOEYQ 94  
 QY 62 SAISGFIKVKYSDEYTFATSNADNHYTMWDDQEVINKASNSKNIRLEKGLYQIKIQYQ 121



QY 445 VRDTGSNWSEVLPQIQTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALK 504  
Db 491 GNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRYAAKDYENPEKTP-PSLTLKDALK 549  
QY 505 IAF--GFNEPNNGLOYQKIDTEPDF--NFDQOTSQNIKNQLAEL-----NATNIYTVLD 555  
Db 549 LSPYDEIKEIEGLLYKKNPIYESSVMYLDNTAKVTKQLNDTTGFKDVSRLYDV-- 606  
QY 556 KIKNAKMNIILIRKRFHYDRNNIAVGADESVMKAEHREVINSTEG-----L 603  
Db 607 --KLTPKMNVTIK-LSILYDN--AESNDNSIGKWTNTNIVSGGNGKKQYSSNNPDANL 660  
QY 604 LNLID-----KDIRKILSGYVIEIDTE-----GLKEVINDRYDMLNI-- 641  
Db 661 TLNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGIEYPIITTKVNVNKNYKELDIIA 720  
QY 642 SSLRODGKTFIDFKYNDKLPYISNPNKYVAVTKENTINPSENGDT-STNGIK-- 698  
Db 721 HNICKNPISIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSEIKQIYSRYGKLE 776  
QY 699 -KILIFSKGYEIG 711  
Db 777 DGLIDKKGGIHYG 790

RESULT 12  
US-09-307-106-8  
Sequence 8, Application US/09307106  
Patent No. 6603063  
GENERAL INFORMATION:  
APPLICANT: Fetteison, Jerald S.  
APPLICANT: Schnepf, H. Ernest  
APPLICANT: Narva, Kenneth E.  
APPLICANT: Stockhoff, Brian A.  
APPLICANT: Schmeits, James  
APPLICANT: Loewer, David  
APPLICANT: Dullum, Charles Joseph  
APPLICANT: Muller-Cohn, Judy  
APPLICANT: Stamp, Lisa  
APPLICANT: Morrill, George  
APPLICANT: Finstad-Lee, Stacey  
TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide  
TITLE OF INVENTION: Sequences Which Encode These Toxins  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/307,106  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,848  
FILING DATE: 30-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/960,780  
FILING DATE: 30-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/073,898  
FILING DATE: 05-MAY-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 881 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE: PS177C8a  
INDIVIDUAL ISOLATE: PS177C8a  
US-09-307-106-8

Query Match 20.8%; Score 759; DB 4; Length 881;  
Best Local Similarity 29.7%; Pred. No. 3.5e-48;  
Matches 236; Conservative 136; Mismatches 294; Indels 128; Gaps 32;

QY 4 QENRLNSESSESSQGLLYFSDLNTOAPMVVTSSTGDLSPSSLEN--IFSENQYFO 61  
Db 39 QKQK--QKEMDRKGLLYFKGKDF-SNLTWPAFTRDSTLIYDQQTANKLLDKQOEYQ 94  
QY 62 SAIWSGFIKVKSDDEVTEATSADNHVTWVDDOEIVNKASNSNKIRLEKRLYQIKIYQ 121  
Db 95 SIRWIGLIQSKETGDFTFNLSEDEQALIBINGKLIISNKGKQVHLEKGLVPKIEYQ 154  
QY 122 RENPTEKGLD-----FKLYWDSQNKKEVISSDNLQLPKOKSS-----N 162  
Db 155 SD--TKFNIDSKTFKELKLFKIDSONQVQVQDELNPFENFKKESQEFLLAKPSKINLFT 212  
QY 163 SRKRSSTAGTVPDRDNDGIPDSLEVEGYTVQVKNRTFLSPWISNIHEKGLTKYKSS 222  
Db 213 QKMKREIDED---TDIDGDSIPDLWEENGTYI---QNRIVAKWDDSL-ASKGYTFVSN 264  
QY 223 PEKWTASDPYSPPEKVTGRIDKNVSPARHPHVAAYPIVHVDMENIILSKNDQSTQNT 282  
Db 265 PLESHTVGDPTDYKKAARDLDLSNAKETENPLVAAFPSPVNVSMKVLSPNENLS--- 320  
QY 283 DSETETISKNTSRTHTSNSS-----TVAIDHLSLSLAGERTWAETMG- 327  
Db 321 ---NSVESSTNTWSVTNTEGASVEAGIGPKGISFQSVNYQHSVEVAQE--WGTSTGN 374  
QY 328 ---LNTADTARLANIRYVNTGTAPIYVNLPTTSLVGLKNGQTATIKADENQISQILAPN 384  
Db 375 TSQENTASAGYLNANRYNVNVTGAIYDKVPTTSFVL--NNDTIATITAKSNSTALNISPG 433  
QY 385 NYFSPKMLAPALNAQKQDASPTITWNYNQFLELEKTKQLRLDQVYGNIAVYFNGR 444  
Db 434 ESYFKGQNGIAITSMDDFNSHPITLNNKQVDNLLNNKPMWLENTQDTG---VYKIDTH 490  
QY 445 VRDTGSNWSEVLPQIQTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALK 504  
Db 491 GNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRYAAKDYENPEKTP-PSLTLKDALK 548  
QY 505 IAF--GFNEPNNGLOYQKIDTEPDF--NFDQOTSQNIKNQLAEL-----NATNIYTVLD 555  
Db 549 LSPYDEIKEIEGLLYKKNPIYESSVMYLDNTAKVTKQLNDTTGFKDVSRLYDV-- 606  
QY 556 KIKNAKMNIILIRKRFHYDRNNIAVGADESVMKAEHREVINSTEG-----L 603  
Db 607 --KLTPKMNVTIK-LSILYDN--AESNDNSIGKWTNTNIVSGGNGKKQYSSNNPDANL 660  
QY 604 LNLID-----KDIRKILSGYVIEIDTE-----GLKEVINDRYDMLNI-- 641  
Db 661 TLNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGIEYPIITTKVNVNKNYKELDIIA 720  
QY 642 SSLRODGKTFIDFKYNDKLPYISNPNKYVAVTKENTINPSENGDT-STNGIK-- 698  
Db 721 HNICKNPISIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSEIKQIYSRYGKLE 776  
QY 699 -KILIFSKGYEIG 711

Db 777 DGLIDKKGIIHYG 790

RESULT 13  
US-09-850-351A-32  
Sequence 32, Application US/09850351A  
Patent No. 6656908  
GENERAL INFORMATION:  
APPLICANT: Feltelson, Gerald S.  
Schnepp, H. Ernest  
Narva, Kenneth B.  
Stockhoff, Brian A.  
Schmeits, James  
Loewer, David  
Dullum, Charles Joseph  
Muller-Cohn, Judy  
Stamp, Lisa  
Morrill, George  
TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide  
Sequences Which Encode These Toxins  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/850,351A  
FILING DATE: 07-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/073,898  
FILING DATE: 06-MAY-1998  
APPLICATION NUMBER: US 08/960,780  
FILING DATE: 30-OCT-1997  
APPLICATION NUMBER: US 60/029,848  
FILING DATE: 30-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-708CD1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 881 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: PS177C8  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-09-850-351A-32  
Query Match 20.8%; Score 759; DB 4; Length 881;  
Best Local Similarity 29.7%; Pred. No. 3.5e-48;  
Matches 236; Conservative 136; Mismatches 294; Indels 128; Gaps 32;  
QY 4 QENLLNESSSQGLCYFSDLNFPQPMVVTSTGDLISPSSELEN--IPSENYQYQ 61  
Db 39 QKNQ---QKENDRKGLLYGFKGDF-SNLTMFAPTRDSTLIYDQQTANKLLDKKQEQY 94  
QY 62 SAIVSGFKVKVKSDEYFATSDADNHVTWVDQVINKASNSKIRLEKRLYQIKYQY 121

95 SIRWIGLIQSKETGDTFNLSEDEQAIIIEINGKIIISNKGKQVHVLEKGLVPIKIEYQ 154  
122 RENPTEKGLD-----FKLYWTDSONKKEVIGSDNLQLPELKQKSS-----N 162  
155 SD--TKFNIDSKTFKELKLFKIDSQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 212  
163 SRKRSSTAGPTVPDRNDGIPDSLEVBGYTVVQVQVQVQVQVQVQVQVQVQVQV 222  
213 QKMKREIDED--TDTGDSIPDLMEENGYYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264  
223 PEKWTASDPYSDPEKVTGRIDKQVSPPEARHPLVAAYPIVHVDMENIILSKNEDOSTQNT 282  
265 PLESHTVGDPYTDYERKAARDLDSNAKETFNPLVAAPFVNVSMKVVILSPENLS---- 320  
283 DSETRTISKNSTSTRTHTSNNS-----TVAIDHSLSLAGERTWAETMG- 327  
321 ---NSVESHSSTNWSYTNTEGASVEAGIGPKGISFGVSNVYQHSSTVAQE--WGTSTGN 374  
328 ---LNTADTARLANIRVYNTGTAPYVNVLPPTSLVGLKNOTLATIKADENQLSOLAPN 384  
375 TSQFNTASAGYLNANRVNNVGTGALYDVKPTTISFVL-NDIITATITAKSNSTALISPG 433  
385 NYTPSKNLAPIALNAQKASSTPTITWYVQVLEKTKQLRLDQVYVNIATYFENG 444  
434 ESYPKKGQNGIAITSMDDFNSHEITLKKQVDNLLANKPMLENTQTDG---VYKIDTH 490  
445 VRVDTGSNWSEVLPOIQTETARIIFNGKOLNVERBIAAVNPSPDLETTKPDMLKEALK 504  
491 GNVITGGEWGVIOQIKAKTASIIIVDGE-RVAEKVRAAKDYENPEDKT-PSLTLDKALK 548  
505 IAF--GFNEPENGVLQVQKDIETBDF--NPDQOOTSQNIKNQLAEL-----NATNIYTVLD 555  
549 LSYPDKEIKETGLLYKKNKPIYESSVNTYLDENTAKEVTKQLNDTTGKFKDVSHLYDV-- 606  
556 KIKUNAKONLIRDKRFHYDRNNIANGADESVVKEAHRREVINSSSTEG-----L 603  
607 --KLTPKQNVTK-LSILYDN---AESNDNSICKWNTNIVSGNNGKKQYSSNPDANL 660  
604 LLNID-----KDIRKILSGVIVEIETE-----GLKEVINDRYDMLNI-- 641  
661 TLNTDAQEKLNKRDYVILYMKSEKNTQCEIIDGIEIYITTKTVNVNKNKDYKRLDIIA 720  
642 SSLRQDQKTFIDFKKYNKXKPLIYSNPNYKVNVAVTKEINTIINPSENGDT-STNGIK-- 698  
721 HNIKSNPISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSSBIKQIYSRYGIKLE 776

Db 699 -KILIFSKGYEIG 711  
777 DGLIDKKGIIHYG 790

RESULT 14  
US-08-471-033-5  
Sequence 5, Application US/08471033  
Patent No. 5770696  
GENERAL INFORMATION:  
APPLICANT: Warren, Gregory W  
APPLICANT: Koziel, Michael G  
APPLICANT: Mullins, Martha A  
APPLICANT: Nye, Gordon J  
APPLICANT: Carr, Brian  
APPLICANT: Desai, Nalini M  
APPLICANT: Kostichka, N. Kristy  
APPLICANT: Duck, Nicholas B  
APPLICANT: Estruch, Juan J  
TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA





; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 884 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-044-5

Query Match 20.8%; Score 757; DB 2; Length 884;
Best Local Similarity 29.7%; Pred. No. 5e-48;
Matches 238; Conservative 131; Mismatches 289; Indels 144; Gaps 32;

QY 4 QENRLINESSESSQGLGYFFSGLNFQAPMVVTSSTTGDLSIPSSSELEN--IPSENQYFQ 61
Db :
42 QKNQ--QKEMDRKGLGYFPKGD--SNLTWFAPTRDSTLIYDQOTANKLLDKKQBEYQ 97
QY :
62 SAIWGFIKVKSDSYTATSDNHTVMVDQEVINKASNSNKIRLEKGRLYQIKIQYQ 121
Db :
98 SIWIGLIQSKETGDTFNLSEDEQAIIIEINGKIIISNKGKEKQVVHLEKGLVPIKIEYQ 157
QY :
122 RENPTEKGLD-----FKLYWTDSONKKEVISNDNLQLPKQKSS-----N 162
Db :
158 SD--TKFNIDSNTFKELKLFKIDSQNPQOQVQDELNPFPNKESQEFLLAKPSKINLFT 215
QY :
163 SRKKSTSGAPVPRDNDGIPDSLEVBGYTVDVQNKRTFLSPWISNHEKKGLTKYKSS 222
Db :
216 QMKKEIDED--TPTDGSIPDLWEENGYYT-----QNRIVAKWDDSL-ASKGYTKFVSN 267
QY :
223 PEKMWSTADPYGDFEKVTRGRIDKQVSPEARHPLVAAPYIVHVDVMEIILSKNEDOSTQNT 282
Db :
268 PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAAPFVSVNVMKVLSPNENLS---- 323
QY :
283 DSETRTISKNTSRTSTNSNSS-----TVAIDHLSLAGERTWAETWG- 327
Db :
324 ---NSVESHSSTNWSYTTTEGASVEAGIGPKGISFGVSVNYQHSSETVAQE--WGTSTGN 377
QY :
328 ---LNTADTARLNANRYVNTGTAPIYVNLPTTSLVLGKNOFLATIKADENQLSOILAPN 384
Db :
378 TSQFNTASAGYLNANRYVNVGTGAIYDVKPTTSFVL--NNDTIAITAKSNSTALNISPG 436
QY :
385 NYYPKXNLAPIANAQKADASSPTTNVYNQFLEKTKQLRLDPTDQVYGNATYNFENGR 444
Db :
437 ESYFKEGQGGIAITSDDFNSHPITLKKQVDNLLNNKPMLETNQTDG---VYKIKDTH 493
QY :
445 VRYDTGSNKSEVLPOIQTETARTIIFNGKDLNLVERRIAAVNPSPDPLETTKPDMTLKEALK 504
Db :
494 GNVITGGEWNGVTQQIKAKTASLIYDDGE--RVAEKRAAKOYENPEDKT--PSLTUKDAK 551
QY :
505 IAF--GFNEPENGLOYQKDIETEDF--NFDQOTSQNIKNQLAEL-----NATNIYTVLD 555
Db :
552 LSYPDKEIKETEGLLYYKKNPIYESSVMTYLDENTAKEVTQLND\*TKGFKDVSHLYDV-- 609
QY :
556 KIKLNAMNILLRDKPFDNRNIAVGADESVVVKEAHREVINSSTEG-----L 603
Db :
610 --KLTPKMNVTIK--LSILYDN---ABSNDNSIGKWTNTNIVSGGNGKQYSSNNPDANL 663
QY :
604 LLNID-----KDIRKILSGYIVEIEDTE-----GLKEVINDRYDMLN--- 640
Db :
664 TLNTDAQELKNKRDYVYISLYMKSEKNTQCEITIDGEIYPIITTKVTVNWKDNYKELDIIA 723
QY :
641 -----ISSLRQGGKTFDFKKYNDKPLYSNPYKNVYAVTKENTINPSENGDT- 692
Db :
724 HNTKSNPISSLH-----IKTNDIEITLFWDDISI--TDVASIKPEN--LTDSEIKQIY 771
QY :
693 STNGIK---KILIFSKKGYEIG 711
Db :
772 SRYGKLEGGILLIDKKGIIHYG 793

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 35.4429 Seconds  
(without alignments)  
5560.545 Million cell updates/sec

Title: US-09-848-909A-16

Perfect score: 3648  
Sequence: 1 EVKQENRLNSESSESSOGLL.....TSTNGIKILIFSKGYEIG 711

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/2/pubpaa/US07\_PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/2/pubpaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3607	98.9	735	15	US-10-410-647-30
2	3607	98.9	736	12	US-09-848-909-1
3	3607	98.9	736	12	US-09-848-909-2
4	3607	98.9	736	12	US-09-848-909-3
5	3607	98.9	736	12	US-09-848-909-4
6	3607	98.9	736	12	US-09-848-909-5
7	3607	98.9	736	12	US-09-848-909-6
8	3607	98.9	736	12	US-09-848-909-7
9	3607	98.9	736	12	US-09-848-909-8
10	3607	98.9	736	12	US-09-848-909-9
11	3607	98.9	736	12	US-09-848-909-10
12	3607	98.9	736	12	US-09-848-909-11
13	3607	98.9	736	12	US-09-848-909-12
14	3607	98.9	736	12	US-09-848-909-13
15	3607	98.9	736	12	US-09-848-909-14

16	3607	98.9	736	12	US-09-848-909-15	Sequence 15, Appl
17	3607	98.9	736	12	US-09-848-909-16	Sequence 16, Appl
18	3607	98.9	736	12	US-09-848-909-17	Sequence 17, Appl
19	3607	98.9	736	12	US-09-848-909-18	Sequence 18, Appl
20	3607	98.9	736	12	US-09-848-909-19	Sequence 19, Appl
21	3607	98.9	736	12	US-09-848-909-20	Sequence 20, Appl
22	3607	98.9	736	12	US-09-848-909-21	Sequence 21, Appl
23	3607	98.9	736	12	US-09-848-909-22	Sequence 22, Appl
24	3607	98.9	736	15	US-10-442-502-7	Sequence 7, Appl
25	3607	98.9	763	15	US-10-442-502-5	Sequence 5, Appl
26	3607	98.9	764	15	US-10-442-502-6	Sequence 6, Appl
27	3604	98.8	735	12	US-10-402-466A-9	Sequence 9, Appl
28	3604	98.8	735	12	US-10-402-466A-13	Sequence 13, Appl
29	3604	98.8	735	14	US-10-332-282-13	Sequence 13, Appl
30	3600	98.7	735	12	US-09-848-909-30	Sequence 30, Appl
31	3600	98.7	764	12	US-10-253-286-681	Sequence 681, App
32	3600	98.7	764	15	US-10-245-871-681	Sequence 681, App
33	3586	98.3	764	9	US-09-747-521-4	Sequence 4, Appl
34	3586	98.3	764	13	US-10-106-014-4	Sequence 4, Appl
35	3586	98.3	764	13	US-10-105-695-4	Sequence 4, Appl
36	3586	98.3	764	14	US-10-105-694-4	Sequence 4, Appl
37	2896	79.2	599	12	US-09-848-909-24	Sequence 24, Appl
38	2891	79.2	595	14	US-10-332-282-11	Sequence 11, Appl
39	2764	75.8	573	12	US-10-402-466A-22	Sequence 22, Appl
40	2752	75.4	569	15	US-10-442-502-8	Sequence 8, Appl
41	2342	64.2	487	14	US-10-332-282-5	Sequence 5, Appl
42	2077	56.9	423	12	US-10-402-466A-24	Sequence 24, Appl
43	2027	55.6	426	14	US-10-332-282-9	Sequence 9, Appl
44	1478	40.5	318	14	US-10-332-282-7	Sequence 7, Appl
45	1338	36.7	258	14	US-10-332-282-3	Sequence 3, Appl

## ALIGNMENTS

## RESULT 1

US-10-410-647-30  
Sequence 30, Application US/10410647  
Publication No. US20030235818A1  
GENERAL INFORMATION:  
APPLICANT: PLEXUS VACCINE, INC.  
APPLICANT: Katritch, Vsevolod  
APPLICANT: Bordner, Andrew  
APPLICANT: Deans, Robert  
APPLICANT: Summer, Mary  
TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME  
FILE REFERENCE: PLEX1110-1  
CURRENT APPLICATION NUMBER: US/10/410,647  
CURRENT FILING DATE: 2003-04-08  
PRIOR APPLICATION NUMBER: US 60/373,668  
PRIOR FILING DATE: 2002-04-17  
PRIOR APPLICATION NUMBER: US 60/371,256  
PRIOR FILING DATE: 2002-04-08  
PRIOR APPLICATION NUMBER: US 60/371,250  
PRIOR FILING DATE: 2002-04-08  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 30  
LENGTH: 735  
TYPE: PRT  
ORGANISM: Bacillus anthracis  
US-10-410-647-30

Query Match 98.9%; Score 3607; DB 15; Length 735;

Best Local Similarity 96.3%; Pred. No. 1.2e-262;

Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSOGLLYEFDLNFQAPVMVTSITGDLISPSSELENIPSENYQF 60

DB 1 EVKQENRLNSESSESSOGLLYEFDLNFQAPVMVTSITGDLISPSSELENIPSENYQF 60

QY 61 QSAWSGFIKYKSGDEYTFATSDAHVMTWYDDQEVINKASNSKIRLEKRLYQIKIQY 120

```

Db 61 QSAIWSGFIKVKSDYTFATSDADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNRKRGSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNRKRGSTAGPTVPDRDN 180
Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 -----SNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASFPIIGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA 420
Qy 397 LNAQKASSTPTIMYNOQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 456
Db 421 LNAQDFFSSTPTIMYNOQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Qy 457 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 516
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540
Qy 517 QYQKDIETFDNFDOQTSONIKQLAELNATNIYVLDKIKLAKQNNILIRDKRFHYDR 576
Db 541 QYQKDIETFDNFDOQTSONIKQLAELNATNIYVLDKIKLAKQNNILIRDKRFHYDR 600
Qy 577 NNIAVGADESVMKEAHREVNSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 636
Db 601 NNIAVGADESVMKEAHREVNSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 637 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNVYAVTKENTIINPSENGDSTNG 696
Db 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNVYAVTKENTIINPSENGDSTNG 720
Qy 697 IKKILIFSCKGYEIG 711
Db 721 IKKILIFSCKGYEIG 735

```

## RESULT 2

```

US-09-848-909-1
; Sequence 1, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-1

```

```

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

```

```

Qy 1 EVKQENRLNSESSESSQGLGYFSDLANFQAPMVVTSSTTGDLSIPSSSELENIIPSENOVF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLANFQAPMVVTSSTTGDLSIPSSSELENIIPSENOVF 60
Qy 61 QSAIWSGFIKVKSDYTFATSDADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Db 61 QSAIWSGFIKVKSDYTFATSDADNHVMTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNRKRGSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNRKRGSTAGPTVPDRDN 180
Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 -----SNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASFPIIGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYYPKSLAPIA 420
Qy 397 LNAQKASSTPTIMYNOQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 456
Db 421 LNAQDFFSSTPTIMYNOQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
Qy 457 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 516
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540
Qy 517 QYQKDIETFDNFDOQTSONIKQLAELNATNIYVLDKIKLAKQNNILIRDKRFHYDR 576
Db 541 QYQKDIETFDNFDOQTSONIKQLAELNATNIYVLDKIKLAKQNNILIRDKRFHYDR 600
Qy 577 NNIAVGADESVMKEAHREVNSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 636
Db 601 NNIAVGADESVMKEAHREVNSTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 637 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNVYAVTKENTIINPSENGDSTNG 696
Db 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNVYAVTKENTIINPSENGDSTNG 720
Qy 697 IKKILIFSCKGYEIG 711
Db 721 IKKILIFSCKGYEIG 735

```

## RESULT 3

```

US-09-848-909-2
; Sequence 2, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis

```

US-09-848-909-2

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTTGDLSIPSELENIPISENQYF 60  
Db 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTTGDLSIPSELENIPISENQYF 60  
Qy 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPELKQKSSNRKGRSTASGTPVDPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPELKQKSSNRKGRSTASGTPVDPDRDN 180  
Qy 181 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 181 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
Qy 397 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGNSWSEV 456  
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGNSWSEV 480  
Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKALKIAPGFNEPNGNL 516  
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKALKIAPGFNEPNGNL 540  
Qy 517 QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKNAKMNLIRDKRFHYDR 576  
Db 541 QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKNAKMNLIRDKRFHYDR 600  
Qy 577 NNIAVGADESUVKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 636  
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Qy 637 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVVYAVTKENTIINPSENGDTSTNG 696  
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720  
Qy 697 IKKILIFSKKGYEIG 711  
Db 721 IKKILIFSKKGYEIG 735

RESULT 4

US-09-848-909-3  
; Sequence 3, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848, 909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: Fast-Seq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-3

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTTGDLSIPSELENIPISENQYF 60  
Db 1 EVKQENRLNESSSQGLLYYFSDLPQAPMVVTSSTTGDLSIPSELENIPISENQYF 60  
Qy 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPELKQKSSNRKGRSTASGTPVDPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEV:SSDNQLPELKQKSSNRKGRSTASGTPVDPDRDN 180  
Qy 181 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 181 QSAIWSGFIKVKKSDEYTPATSDADNHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120  
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 420  
Qy 397 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGNSWSEV 456  
Db 421 LNAQKDASSPTITMNYNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGNSWSEV 480  
Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKALKIAPGFNEPNGNL 516  
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKALKIAPGFNEPNGNL 540  
Qy 517 QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKNAKMNLIRDKRFHYDR 576  
Db 541 QYQCKDITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKNAKMNLIRDKRFHYDR 600  
Qy 577 NNIAVGADESUVKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 636  
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Qy 637 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVVYAVTKENTIINPSENGDTSTNG 696  
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720  
Qy 697 IKKILIFSKKGYEIG 711  
Db 721 IKKILIFSKKGYEIG 735

RESULT 5

US-09-848-909-4  
; Sequence 4, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 736  
; TYPE: PRF  
; ORGANISM: Bacillus anthracis  
US-09-848-909-4

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
QY 61 QSAIWSGFTKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFTKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLPQPELKQSSNSRKRSTSGAPVDPDRN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLPQPELKQSSNSRKRSTSGAPVDPDRN 180  
QY 181 DGPDSLEVEGYTVDKNRTFLSPWISNIHEKGLTKYKSPKSWSTASDYPDFEKT 240  
DB 181 DGPDSLEVEGYTVDKNRTFLSPWISNIHEKGLTKYKSPKSWSTASDYPDFEKT 240  
QY 241 GRIDKNVSPKARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSTRITSKNTSRTHT 300  
DB 241 GRIDKNVSPKARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSTRITSKNTSRTHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
DB 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 456  
DB 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 457 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 516  
DB 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540  
QY 517 QYQKDIITEFDNFDDQTSQNTKNQLAELNATNITYVLDKIKLNAKMILLRDKRFHYDR 576  
DB 541 QYQKDIITEFDNFDDQTSQNTKNQLAELNATNITYVLDKIKLNAKMILLRDKRFHYDR 600  
QY 577 NNIAVGADESUVKEAREVINSSTEGLLNIDKDKRKILSGYVIEIDTEGLKEVINDRY 636  
DB 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKDKRKILSGYVIEIDTEGLKEVINDRY 660  
QY 637 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVKENTIINPSENGDSTNG 696  
DB 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVKENTIINPSENGDSTNG 720  
QY 697 IKKILFSKGYEIG 711  
DB 721 IKKILFSKGYEIG 735

RESULT 6  
US-09-848-909-5

; Sequence 5, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 736  
; TYPE: PRF  
; ORGANISM: Bacillus anthracis  
US-09-848-909-5

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
QY 61 QSAIWSGFTKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFTKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLPQPELKQSSNSRKRSTSGAPVDPDRN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLPQPELKQSSNSRKRSTSGAPVDPDRN 180  
QY 181 DGPDSLEVEGYTVDKNRTFLSPWISNIHEKGLTKYKSPKSWSTASDYPDFEKT 240  
DB 181 DGPDSLEVEGYTVDKNRTFLSPWISNIHEKGLTKYKSPKSWSTASDYPDFEKT 240  
QY 241 GRIDKNVSPKARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSTRITSKNTSRTHT 300  
DB 241 GRIDKNVSPKARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSTRITSKNTSRTHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
DB 361 NANIRVYNTGTAPIYVNLPTTSLVKGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 456  
DB 421 LNAQDDFSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 457 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 516  
DB 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540  
QY 517 QYQKDIITEFDNFDDQTSQNTKNQLAELNATNITYVLDKIKLNAKMILLRDKRFHYDR 576  
DB 541 QYQKDIITEFDNFDDQTSQNTKNQLAELNATNITYVLDKIKLNAKMILLRDKRFHYDR 600  
QY 577 NNIAVGADESUVKEAREVINSSTEGLLNIDKDKRKILSGYVIEIDTEGLKEVINDRY 636  
DB 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKDKRKILSGYVIEIDTEGLKEVINDRY 660  
QY 637 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVKENTIINPSENGDSTNG 696  
DB 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVKENTIINPSENGDSTNG 720  
QY 697 IKKILFSKGYEIG 711

Db 721 IKKILIFSKGYEIG 735

## RESULT 7

US-09-848-909-6

; Sequence 6, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-6

Query Match 98.9%; Score 3607; DB 12; Length 736;

Best Local Similarity 96.3%; Pred. No. 1.2e-262;

Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60

Qy 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120

Db 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240

Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240

Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336

Db 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 337 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKKNLAPTA 396

Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKKNLAPTA 420

Qy 397 LNAQKDASSPTITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVVRVDTGSNWSEV 456

Db 421 LNAQDDFSSPTITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVVRVDTGSNWSEV 480

Qy 457 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKEALKIAGFNPENGNL 516

Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKEALKIAGFNPENGNL 540

Qy 517 QYQKDIPTDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNTLIRDKRPHYDR 576

Db 541 QYQKDIPTDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNTLIRDKRPHYDR 600

Qy 577 NNIAVGADESIVKAEHREVNNSSTEGLLNIDKDKIRKILSGYIVETEDTEGLKEVINDRY 636

Db 601 NNIAVGADESIVKAEHREVNNSSTEGLLNIDKDKIRKILSGYIVETEDTEGLKEVINDRY 660

Qy 637 DMLNITSSLRQDKTIDFKKYNDKLPYISNPNYKNTYAVTKENTIIINPSENGDTSTNG 696

Db 661 DMLNITSSLRQDKTIDFKKYNDKLPYISNPNYKNTYAVTKENTIIINPSENGDTSTNG 720

Qy 697 IKKILIFSKGYEIG 711

Db 721 IKKILIFSKGYEIG 735

## RESULT 8

US-09-848-909-7

; Sequence 7, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-7

Query Match 98.9%; Score 3607; DB 12; Length 736;

Best Local Similarity 96.3%; Pred. No. 1.2e-262;

Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60

Qy 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120

Db 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGLYQIKIY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTASDPSDFEKT 180

Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240

Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTASDPSDFEKT 240

Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336

Db 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 337 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKKNLAPTA 396

Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSQILAPNNYPSKKNLAPTA 420

Qy 397 LNAQKDASSPTITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVVRVDTGSNWSEV 456

Db 421 LNAQDDFSSPTITMNNQFLEBKTKQLRLDTPQVYGNITATYNFENGRVVRVDTGSNWSEV 480

Qy 457 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKEALKIAGFNPENGNL 516

Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETKPDMTLKEALKIAGFNPENGNL 540

Qy 517 QYQKDIPTDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNTLIRDKRPHYDR 576

Db 541 QYQKDIETFDNFQOQTSQNIKNQLAELNATNIVTLDKIKLNKAKNILLRDKRPHYDR 600  
QY 577 NNIAGADESVVKEAHREVINSSTGLLLNIDKDIRKILSGYIVIEIETBGLKEVINDRY 636  
Db 601 NNIAGADESVVKEAHREVINSSTGLLLNIDKDIRKILSGYIVIEIETBGLKEVINDRY 660  
QY 637 DMLNISSLRQDGKTFIDFKKYNKPLXYISNPNYKVVAVTKENTIIINPSENGDTSTNG 696  
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLXYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720  
QY 697 IKKILIFSKKGYEIG 711  
Db 721 IKKILIFSKKGYEIG 735  
RESULT 9  
US-09-848-909-8  
; Sequence 8, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-8

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMWVTSSTTGDLSIPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMWVTSSTTGDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQSSNSRKRSTSSAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQSSNSRKRSTSSAGTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 336  
Db 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYVLPSTSLVGLKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYVLPSTSLVGLKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKQDASPTITWYNYQFLEKTKQLRLDITDQVYGNATYNFENGRVVRVDTGNSWSEV 456  
Db 421 LNAQKQDASPTITWYNYQFLEKTKQLRLDITDQVYGNATYNFENGRVVRVDTGNSWSEV 480

QY 457 LPOIQETTARIIPNGKDLNLVERRIAANVPSDPLETTTKDPMTLKALKTAFGNPNGNL 516  
Db 481 LPOIQETTARIIPNGKDLNLVERRIAANVPSDPLETTTKDPMTLKALKTAFGNPNGNL 540  
QY 517 QYQKDIETFDNFQOQTSQNIKNQLAELNATNIVTLDKIKLNKAKNILLRDKRPHYDR 576  
Db 541 QYQKDIETFDNFQOQTSQNIKNQLAELNATNIVTLDKIKLNKAKNILLRDKRPHYDR 600  
QY 577 NNIAGADESVVKEAHREVINSSTGLLLNIDKDIRKILSGYIVIEIETBGLKEVINDRY 636  
Db 601 NNIAGADESVVKEAHREVINSSTGLLLNIDKDIRKILSGYIVIEIETBGLKEVINDRY 660  
QY 637 DMLNISSLRQDGKTFIDFKKYNKPLXYISNPNYKVVAVTKENTIIINPSENGDTSTNG 696  
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLXYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720  
QY 697 IKKILIFSKKGYEIG 711  
Db 721 IKKILIFSKKGYEIG 735  
RESULT 10  
US-09-848-909-9  
; Sequence 9, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-9

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMWVTSSTTGDLSIPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMWVTSSTTGDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQSSNSRKRSTSSAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQPELKQSSNSRKRSTSSAGTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 336  
Db 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERWAEWTMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYVLPSTSLVGLKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396



```
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASSTPTIMYVNTQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 456
Db 421 LNAQDDFSTPTIMYVNTQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 480
Qy 457 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 516
Db 481 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 540
Qy 517 QYQGGKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 576
Db 541 QYQGGKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
Qy 577 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 637 DMLNLSLRQDGKTFIDFKYNDKULPLYISNPYKVNVAATYKENTIIINPSENGDTSTNG 696
Db 661 DMLNLSLRQDGKTFIDFKYNDKULPLYISNPYKVNVAATYKENTIIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

RESULT 11
US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-10

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDYEYFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYEYFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSQNKKEVSSDNQLPELKQKSNRSRKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSQNKKEVSSDNQLPELKQKSNRSRKRSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKVT 240
Db 181 DGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKVT 240
Qy 241 GRIDKXVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKXVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
```

```
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHAFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKQDASSTPTIMYVNTQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 456
Db 421 LNAQDDFSTPTIMYVNTQFLELEKTKQLRLDQVYGNIAATYFNGRVRVDTGNSWSEV 480
Qy 457 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 516
Db 481 LPOIQTETARIIFNGKDLNVLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPNGNL 540
Qy 517 QYQGGKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 576
Db 541 QYQGGKDIETFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKMNILIRDKRPHYDR 600
Qy 577 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 636
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Qy 637 DMLNLSLRQDGKTFIDFKYNDKULPLYISNPYKVNVAATYKENTIIINPSENGDTSTNG 696
Db 661 DMLNLSLRQDGKTFIDFKYNDKULPLYISNPYKVNVAATYKENTIIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735

RESULT 12
US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-11

Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDYEYFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYEYFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Qy 121 QRENTEKGLDFKLYWTDSQNKKEVSSDNQLPELKQKSNRSRKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSQNKKEVSSDNQLPELKQKSNRSRKRSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKVT 240
```

Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASTASDPYDFEKT 240  
QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVAHAFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYNNVPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYNNVPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASSTPTIMYNNQFLELEKTKQILRLDQVYGNIAATYNFENGRVRYDVTGSNWSEV 456  
Db 421 LNAQDDFSSPTIMYNNQFLELEKTKQILRLDQVYGNIAATYNFENGRVRYDVTGSNWSEV 480  
QY 457 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 516  
Db 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540  
QY 517 QYQKDIITDFDNDQSTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 576  
Db 541 QYQKDIITDFDNDQSTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600  
QY 577 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 636  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 637 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 696  
Db 661 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 697 IKKILIFSKGYEIG 711  
Db 721 IKKILIFSKGYEIG 735

## RESULT 13

US-09-848-909-12  
; Sequence 12, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:

; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-12

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESQGLGYYFSDLNFPQAPMVVTSSTGDLSPSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLGYYFSDLNFPQAPMVVTSSTGDLSPSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDBYTTFATSDADNHVTWDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDBYTTFATSDADNHVTWDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKCRSTASGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKCRSTASGTPVDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASTASDPYDFEKT 240  
QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVAHAFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 337 NANIRYVNTGTAPIYNNVPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
Db 361 NANIRYVNTGTAPIYNNVPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 397 LNAQKDASSTPTIMYNNQFLELEKTKQILRLDQVYGNIAATYNFENGRVRYDVTGSNWSEV 456  
Db 421 LNAQDDFSSPTIMYNNQFLELEKTKQILRLDQVYGNIAATYNFENGRVRYDVTGSNWSEV 480  
QY 457 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 516  
Db 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPNGNL 540  
QY 517 QYQKDIITDFDNDQSTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 576  
Db 541 QYQKDIITDFDNDQSTQNT:KNQLAELNATNIYTVLDKIKLNAMNIIIRDKRFHYDR 600  
QY 577 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 636  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 637 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 696  
Db 661 DMLNISLRODQKTFIDFKKNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 697 IKKILIFSKGYEIG 711  
Db 721 IKKILIFSKGYEIG 735

## RESULT 14

US-09-848-909-13  
; Sequence 13, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:

; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-13

Query Match 98.9%; Score 3607; DB 12; Length 736;  
Best Local Similarity 96.3%; Pred. No. 1.2e-262;  
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESQGLGYYFSDLNFPQAPMVVTSSTGDLSPSELENIPSENOYF 60

```
Db 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIOY 120
Db 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIOY 120
Qy 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPELKQKSSNSRKRSTSGPTVPRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPELKQKSSNSRKRSTSGPTVPRDN 180
Qy 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKVT 240
Db 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKDSSTPTIMVYNQFLEKTKQLRDLTDQVYGNATYVNFENGRVVRVDTGSNWSEV 456
Db 421 LNAQDFFSPTIMVYNQFLEKTKQLRDLTDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGL 516
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGL 540
Qy 517 QYQKDIETEDFDFNDOOTSQNKQLAELNATNIYVLDKIKLAKXNMLIRDKRFHYDR 576
Db 541 QYQKDIETEDFDFNDOOTSQNKQLAELNATNIYVLDKIKLAKXNMLIRDKRFHYDR 600
Qy 577 NNAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 636
Db 601 NNAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 637 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVTAKENTIINPSENGDTSTNG 696
Db 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVTAKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735
```

## RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-14
```

```
Query Match 98.9%; Score 3607; DB 12; Length 736;
Best Local Similarity 96.3%; Pred. No. 1.2e-262;
Matches 708; Conservative 0; Mismatches 3; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIOY 120
Db 61 QSAIWSGFIKVKKSDYTFATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIOY 120
Qy 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPELKQKSSNSRKRSTSGPTVPRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPELKQKSSNSRKRSTSGPTVPRDN 180
Qy 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKVT 240
Db 181 DGIPODSEVBEGYVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRISKNTSTSRHT 300
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336
Db 301 SEVHGNAEVHASPFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 396
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 397 LNAQKDSSTPTIMVYNQFLEKTKQLRDLTDQVYGNATYVNFENGRVVRVDTGSNWSEV 456
Db 421 LNAQDFFSPTIMVYNQFLEKTKQLRDLTDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGL 516
Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGL 540
Qy 517 QYQKDIETEDFDFNDOOTSQNKQLAELNATNIYVLDKIKLAKXNMLIRDKRFHYDR 576
Db 541 QYQKDIETEDFDFNDOOTSQNKQLAELNATNIYVLDKIKLAKXNMLIRDKRFHYDR 600
Qy 577 NNAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 636
Db 601 NNAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 637 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVTAKENTIINPSENGDTSTNG 696
Db 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVTAKENTIINPSENGDTSTNG 720
Qy 697 IKKILIFSKKGYEIG 711
Db 721 IKKILIFSKKGYEIG 735
```

Search completed: May 3, 2004, 20:02:16  
Job time : 38.4429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 169.814 Seconds  
(without alignments)  
4086.665 Million cell updates/sec

Title: US-09-848-909A-16

Perfect score: 3648

Sequence: 1 EVKQENRLNSESQGL.....TSTNGIKLIPFKKGVEIG 711

Scoring table: BLOSUM62

Gap0 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_Main:

```

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/2/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/2/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/2/paa/US107_COMB.pep.*
33: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
-----				

1	3648	100.0	711	1	PCT-US03-35733-16	Sequence 16, Appl
2	3648	100.0	711	23	US-09-848-909A-16	Sequence 16, Appl
3	3642	99.8	711	1	PCT-US03-35733-15	Sequence 15, Appl
4	3642	99.8	711	23	US-09-848-909A-15	Sequence 15, Appl
5	3635	99.6	711	1	PCT-US03-35733-14	Sequence 14, Appl
6	3635	99.6	711	23	US-09-848-909A-14	Sequence 14, Appl
7	3629	99.5	711	1	PCT-US03-35733-12	Sequence 12, Appl
8	3629	99.5	711	23	US-09-848-909A-12	Sequence 12, Appl
9	3626	99.4	735	1	PCT-US03-35733-13	Sequence 13, Appl
10	3626	99.4	735	23	US-09-848-909A-13	Sequence 13, Appl
11	3620	99.2	735	1	PCT-US03-35733-10	Sequence 10, Appl
12	3620	99.2	735	23	US-09-848-909A-10	Sequence 10, Appl
13	3614	99.1	735	1	PCT-US03-19786-4	Sequence 4, Appl
14	3614	99.1	735	23	US-09-848-909A-2	Sequence 2, Appl
15	3614	99.1	735	23	US-09-848-909A-2	Sequence 2, Appl
16	3613	99.0	735	1	PCT-US03-35733-8	Sequence 8, Appl
17	3613	99.0	735	1	PCT-US03-35733-9	Sequence 9, Appl
18	3613	99.0	735	23	US-09-848-909A-8	Sequence 8, Appl
19	3613	99.0	735	23	US-09-848-909A-8	Sequence 8, Appl
20	3609	98.9	735	1	PCT-US03-35733-7	Sequence 7, Appl
21	3609	98.9	735	1	PCT-US03-35733-23	Sequence 23, Appl
22	3609	98.9	735	23	US-09-848-909A-7	Sequence 7, Appl
23	3609	98.9	735	23	US-09-848-909A-23	Sequence 23, Appl
24	3608	98.9	735	1	PCT-US03-35733-4	Sequence 4, Appl
25	3608	98.9	735	1	PCT-US03-35733-6	Sequence 6, Appl
26	3608	98.9	735	1	PCT-US03-35733-18	Sequence 18, Appl
27	3608	98.9	735	23	US-09-848-909A-4	Sequence 4, Appl
28	3608	98.9	735	23	US-09-848-909A-6	Sequence 6, Appl
29	3608	98.9	735	23	US-09-848-909A-18	Sequence 18, Appl
30	3608	98.9	762	4	US-08-042-745-5	Sequence 5, Appl
31	3607	98.9	735	1	PCT-US03-35733-5	Sequence 5, Appl
32	3607	98.9	735	1	PCT-US03-35733-11	Sequence 11, Appl
33	3607	98.9	735	1	PCT-US03-35733-17	Sequence 17, Appl
34	3607	98.9	735	1	PCT-US03-35733-19	Sequence 19, Appl
35	3607	98.9	735	1	PCT-US03-35733-20	Sequence 20, Appl
36	3607	98.9	735	1	PCT-US03-35733-21	Sequence 21, Appl
37	3607	98.9	735	22	US-09-791-537-43735	Sequence 43735, A
38	3607	98.9	735	23	US-09-848-909A-5	Sequence 5, Appl
39	3607	98.9	735	23	US-09-848-909A-17	Sequence 17, Appl
40	3607	98.9	735	23	US-09-848-909A-11	Sequence 11, Appl
41	3607	98.9	735	23	US-09-848-909A-19	Sequence 19, Appl
42	3607	98.9	735	23	US-09-848-909A-20	Sequence 20, Appl
43	3607	98.9	735	23	US-09-848-909A-21	Sequence 21, Appl
44	3607	98.9	735	30	US-10-410-647-30	Sequence 30, Appl
45	3607	98.9	736	1	PCT-US01-14372A-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
PCT-US03-35733-16  
; Sequence 16, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 711  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-16

Query Match 100.0%; Score 3648; DB 1; Length 711;  
Best Local Similarity 100.0%; Pred. No. 6.8e-287;  
Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPISENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPISENOYF 60
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNLOLPKQKSSNSRKRSTSGAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNLOLPKQKSSNSRKRSTSGAGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SNSNSTVAIDHSLSLAGERTWAEITMGLTADTARLNANIRVNTGTAPIYVNLPTTSLV 360
Db 301 SNSNSTVAIDHSLSLAGERTWAEITMGLTADTARLNANIRVNTGTAPIYVNLPTTSLV 360
QY 361 LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQKQDASSTPTIMYNOFLELEK 420
Db 361 LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQKQDASSTPTIMYNOFLELEK 420
QY 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
Db 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
QY 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNNGNLQYQKDIETFDNFQDQTSQNIKN 540
Db 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNNGNLQYQKDIETFDNFQDQTSQNIKN 540
QY 541 QLAELNATNIYTVLDKIKLNAMNLIIRDRKHYDRNNIAGADESVVKEAHEVINSST 600
Db 541 QLAELNATNIYTVLDKIKLNAMNLIIRDRKHYDRNNIAGADESVVKEAHEVINSST 600
QY 601 EGGLELNDKIRKILSGYIVEIEDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660
Db 601 EGGLELNDKIRKILSGYIVEIEDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660
QY 661 LPLYSNPYKVNVAVTAKNTIINPSENGDSTNGIKKILIFSKKGYEIG 711
Db 661 LPLYSNPYKVNVAVTAKNTIINPSENGDSTNGIKKILIFSKKGYEIG 711

```

## RESULT 2

```

US-09-848-909A-16
; Sequence 16, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-16

```

Query Match 100.0%; Score 3648; DB 23; Length 711;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-287;

```

Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPISENOYF 60
Db 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPISENOYF 60
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNLOLPKQKSSNSRKRSTSGAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVIVSSDNLOLPKQKSSNSRKRSTSGAGTPVDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SNSNSTVAIDHSLSLAGERTWAEITMGLTADTARLNANIRVNTGTAPIYVNLPTTSLV 360
Db 301 SNSNSTVAIDHSLSLAGERTWAEITMGLTADTARLNANIRVNTGTAPIYVNLPTTSLV 360
QY 361 LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQKQDASSTPTIMYNOFLELEK 420
Db 361 LGKQTLATIKADENQSLQILAPNNYPSKNLAPIALNAQKQDASSTPTIMYNOFLELEK 420
QY 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
Db 421 TKQLRLDQVGNATYATNFENGRVVDVTSNWSSEVLPOIQETTARIIFNGKDLNLVERR 480
QY 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNNGNLQYQKDIETFDNFQDQTSQNIKN 540
Db 481 IAAVNPSPLETTKPDMTLKEALKIAFGNPNNGNLQYQKDIETFDNFQDQTSQNIKN 540
QY 541 QLAELNATNIYTVLDKIKLNAMNLIIRDRKHYDRNNIAGADESVVKEAHEVINSST 600
Db 541 QLAELNATNIYTVLDKIKLNAMNLIIRDRKHYDRNNIAGADESVVKEAHEVINSST 600
QY 601 EGGLELNDKIRKILSGYIVEIEDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660
Db 601 EGGLELNDKIRKILSGYIVEIEDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660
QY 661 LPLYSNPYKVNVAVTAKNTIINPSENGDSTNGIKKILIFSKKGYEIG 711
Db 661 LPLYSNPYKVNVAVTAKNTIINPSENGDSTNGIKKILIFSKKGYEIG 711

```

## RESULT 3

```

PCT-US03-35733-15
; Sequence 15, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-15

```

Query Match 99.8%; Score 3642; DB 1; Length 711;  
 Best Local Similarity 99.9%; Pred. No. 2.1e-286;

Matches 710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFFDLNFAQPMVVTSTTGDLSIPSSSELENIIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLLYFFDLNFAQPMVVTSTTGDLSIPSSSELENIIPSENQYF 60

QY 61 QSAIMSGFTKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAIMSGFTKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTDSONKEVISSDNQLQPLKQSSNSRKRSTSGAGTVPDRDN 180  
DB 121 QRENTEKGLDFKLYWTDSONKEVISSDNQLQPLKQSSNSRKRSTSGAGTVPDRDN 180

QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240  
DB 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAEITWGLNTADTARLNANIRYNTGTAPIYVNLPTTSLV 360  
DB 301 SNSNSTVAIDHSLSLAGERTWAEITWGLNTADTARLNANIRYNTGTAPIYVNLPTTSLV 360

QY 361 LGKNQTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIIMYNOFLELEK 420  
DB 361 LGKNQTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIIMYNOFLELEK 420

QY 421 TKQLRLDQVYGNATYNFENGRVVDGTSNWSEVLPOIQTETARIIFNGKDLNLEVER 480  
DB 421 TKQLRLDQVYGNATYNFENGRVVDGTSNWSEVLPOIQTETARIIFNGKDLNLEVER 480

QY 481 IAAVNPSPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540  
DB 481 IAAVNPSPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNAKONTLIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600  
DB 541 QLAELNATNIYTVLDKIKLNAKONTLIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600

QY 601 EGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660  
DB 601 EGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660

QY 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711  
DB 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711

## RESULT 4

US-09-848-909A-15

; Sequence 15, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 711  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-15

Query Match

99.8%; Score 3642; DB 23; Length 711;

Best Local Similarity 99.9%; Pred. No. 2.1e-286;  
Matches 710; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFFDLNFAQPMVVTSTTGDLSIPSSSELENIIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLLYFFDLNFAQPMVVTSTTGDLSIPSSSELENIIPSENQYF 60

QY 61 QSAIMSGFTKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAIMSGFTKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTDSONKEVISSDNQLQPLKQSSNSRKRSTSGAGTVPDRDN 180  
DB 121 QRENTEKGLDFKLYWTDSONKEVISSDNQLQPLKQSSNSRKRSTSGAGTVPDRDN 180

QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240  
DB 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAEITWGLNTADTARLNANIRYNTGTAPIYVNLPTTSLV 360  
DB 301 SNSNSTVAIDHSLSLAGERTWAEITWGLNTADTARLNANIRYNTGTAPIYVNLPTTSLV 360

QY 361 LGKNQTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIIMYNOFLELEK 420  
DB 361 LGKNQTLATIKADENQLSQILAPNNYPSKNLAPIALNAQDASSTPIIMYNOFLELEK 420

QY 421 TKQLRLDQVYGNATYNFENGRVVDGTSNWSEVLPOIQTETARIIFNGKDLNLEVER 480  
DB 421 TKQLRLDQVYGNATYNFENGRVVDGTSNWSEVLPOIQTETARIIFNGKDLNLEVER 480

QY 481 IAAVNPSPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540  
DB 481 IAAVNPSPLETTKPDWTLKEALKIAGFNEPENGLOQYQKDIETEFDFNFQDQTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNAKONTLIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600  
DB 541 QLAELNATNIYTVLDKIKLNAKONTLIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600

QY 601 EGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660  
DB 601 EGLLNIDKDIRKILSGYIVEIETEGLEKEVINDRYDMLNSSLRQDGKTFIDFKKYNDK 660

QY 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711  
DB 661 LPLYISNPYKVNVAVTAKENTIIINPSENGDTSTNGIKKILIFSKKGYEIG 711

## RESULT 5

PCT-US03-35733-14

; Sequence 14, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 711  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-14

Query Match

99.6%; Score 3635; DB 1; Length 711;

```

Best Local Similarity 99.7%; Pred. No. 7.7e-286;
Matches 709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLLYVFDLNFQAPMVTSTTGDLSIPSSSELENIPSENQYF 60
Dd 1 EVKQENRLNESSSQGLLYVFDLNFQAPMVTSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Dd 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASAGTVPDRDN 180
Dd 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASAGTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT 240
Dd 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT 300
Dd 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT 300
Qy 301 SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360
Dd 301 SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360
Qy 361 LGKQNTLATIKADENQSLQILAPNNYPSKNLAPALNAQDASSTPTIMYNOFLELEK 420
Dd 361 LGKQNTLATIKADENQSLQILAPNNYPSKNLAPALNAQDASSTPTIMYNOFLELEK 420
Qy 421 TKQLRLDTPQVGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLVRR 480
Dd 421 TKQLRLDTPQVGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLVRR 480
Qy 481 IAAVNSDPLETTKPDMTLKEALKIAGFNEPENGNILOYGKDIITFDNFDOQTSONIKN 540
Dd 481 IAAVNSDPLETTKPDMTLKEALKIAGFNEPENGNILOYGKDIITFDNFDOQTSONIKN 540
Qy 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600
Dd 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600
Qy 601 EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNK 660
Dd 601 EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNK 660
Qy 661 LPLYISPNYKVNVAVTKEHTIINPSENGDSTNGIKKILIFSKKGYEIG 711
Dd 661 LPLYISPNYKVNVAVTKEHTIINPSENGDSTNGIKKILIFSKKGYEIG 711

```

## RESULT 6

US-09-848-909A-14

```

; Sequence 14. Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-14

```

## RESULT 7

PCT-US03-35733-12

```

; Sequence 12. Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-12

```

```

Query Match 99.6%; Score 3635; DB 23; Length 711;
Best Local Similarity 99.7%; Pred. No. 7.7e-286;
Matches 709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLLYVFDLNFQAPMVTSTTGDLSIPSSSELENIPSENQYF 60
Dd 1 EVKQENRLNESSSQGLLYVFDLNFQAPMVTSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Dd 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASAGTVPDRDN 180
Dd 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASAGTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT 240
Dd 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKXSTASDPYDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT 300
Dd 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTRTHT 300
Qy 301 SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360
Dd 301 SNSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTTSLV 360
Qy 361 LGKQNTLATIKADENQSLQILAPNNYPSKNLAPALNAQDASSTPTIMYNOFLELEK 420
Dd 361 LGKQNTLATIKADENQSLQILAPNNYPSKNLAPALNAQDASSTPTIMYNOFLELEK 420
Qy 421 TKQLRLDTPQVGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLVRR 480
Dd 421 TKQLRLDTPQVGNIAATYFNGRVRVDTGSKNWEVLPOIQTETARIIFNGKDLNLVRR 480
Qy 481 IAAVNSDPLETTKPDMTLKEALKIAGFNEPENGNILOYGKDIITFDNFDOQTSONIKN 540
Dd 481 IAAVNSDPLETTKPDMTLKEALKIAGFNEPENGNILOYGKDIITFDNFDOQTSONIKN 540
Qy 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600
Dd 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAGADESVVKEAHREVINSST 600
Qy 601 EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNK 660
Dd 601 EGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRYDMLNSSLRQDGKTFIDFKKYNK 660
Qy 661 LPLYISPNYKVNVAVTKEHTIINPSENGDSTNGIKKILIFSKKGYEIG 711
Dd 661 LPLYISPNYKVNVAVTKEHTIINPSENGDSTNGIKKILIFSKKGYEIG 711

```

Query Match 99.5%; Score 3629; DB 1; Length 711;  
Best Local Similarity 99.6%; Pred. No. 2.4e-285;  
Matches 708; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFDFLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFDFLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATSAADNHVTMWDDQEVINKASNKIRLEKRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSAADNHVTMWDDQEVINKASNKIRLEKRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKCRSTAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKCRSTAGTPVDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPPTSLV 360  
Db 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPPTSLV 360

QY 361 LGKNOTLATIKADENQSLIAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK 420  
Db 361 LGKNOTLATIKADENQSLIAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK 420

QY 421 TKQLRLDQVYGNATYNFENGRVVDVTGSNWSVLPQIETTARIIFNGKOLNVERR 480  
Db 421 TKQLRLDQVYGNATYNFENGRVVDVTGSNWSVLPQIETTARIIFNGKOLNVERR 480

QY 481 IAAVNSPDLPTTKPDMTLKEALKIAGFNEPENGNYQYQKDIETFDNFDOOTSQNIKN 540  
Db 481 IAAVNSPDLPTTKPDMTLKEALKIAGFNEPENGNYQYQKDIETFDNFDOOTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAVGADESUVKEAHEVINSST 600  
Db 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAVGADESUVKEAHEVINSST 600

QY 601 EGGLELNDKIDKILSGYIIEIEDTEGLKEVINDRYDMLNSSLRODGKTFIDFKKYNK 660  
Db 601 EGGLELNDKIDKILSGYIIEIEDTEGLKEVINDRYDMLNSSLRODGKTFIDFKKYNK 660

QY 661 LPLVSNPNYKVVAVYAVTKENTIINPSENGDTSTNGIKKILIFSCKGYEIG 711  
Db 661 LPLVSNPNYKVVAVYAVTKENTIINPSENGDTSTNGIKKILIFSCKGYEIG 711

RESULT 8  
US-09-848-909A-12  
; Sequence 12, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 711  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-12

Query Match 99.5%; Score 3629; DB 23; Length 711;  
Best Local Similarity 99.6%; Pred. No. 2.4e-285;  
Matches 708; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFDFLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFDFLNFAQPMVVTSTTGDLSIPSELENIPSENYF 60

QY 61 QSAIWSGFIKVKSDSEYTFATSAADNHVTMWDDQEVINKASNKIRLEKRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSAADNHVTMWDDQEVINKASNKIRLEKRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKCRSTAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKCRSTAGTPVDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTADPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPPTSLV 360  
Db 301 SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPIYVLPPTSLV 360

QY 361 LGKNOTLATIKADENQSLIAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK 420  
Db 361 LGKNOTLATIKADENQSLIAPNNYPSKNLAPIALNAQDASSTPTIMYNOFLELEK 420

QY 421 TKQLRLDQVYGNATYNFENGRVVDVTGSNWSVLPQIETTARIIFNGKOLNVERR 480  
Db 421 TKQLRLDQVYGNATYNFENGRVVDVTGSNWSVLPQIETTARIIFNGKOLNVERR 480

QY 481 IAAVNSPDLPTTKPDMTLKEALKIAGFNEPENGNYQYQKDIETFDNFDOOTSQNIKN 540  
Db 481 IAAVNSPDLPTTKPDMTLKEALKIAGFNEPENGNYQYQKDIETFDNFDOOTSQNIKN 540

QY 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAVGADESUVKEAHEVINSST 600  
Db 541 QLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDRNNIAVGADESUVKEAHEVINSST 600

QY 601 EGGLELNDKIDKILSGYIIEIEDTEGLKEVINDRYDMLNSSLRODGKTFIDFKKYNK 660  
Db 601 EGGLELNDKIDKILSGYIIEIEDTEGLKEVINDRYDMLNSSLRODGKTFIDFKKYNK 660

QY 661 LPLVSNPNYKVVAVYAVTKENTIINPSENGDTSTNGIKKILIFSCKGYEIG 711  
Db 661 LPLVSNPNYKVVAVYAVTKENTIINPSENGDTSTNGIKKILIFSCKGYEIG 711

RESULT 9  
PCT-US03-35733-13  
; Sequence 13, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-13



Query Match 99.4%; Score 3626; DB 1; Length 735;  
Best Local Similarity 96.7%; Pred. No. 4.4e-285;  
Matches 711; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 60

QY 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 180  
DB 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 180

QY 181 DOIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
DB 181 DOIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420

QY 397 LNAQKASSTPTIWNVQNFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 456  
DB 421 LNAQKASSTPTIWNVQNFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480

QY 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 516  
DB 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 540

QY 517 QYQKDIETFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAMNILLIRDRPHYDR 576  
DB 541 QYQKDIETFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAMNILLIRDRPHYDR 600

QY 577 NNIAVGADESVMKEAREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 636  
DB 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660

QY 637 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 696  
DB 661 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 720

QY 697 IKKILIFSKKGYEIG 711  
DB 721 IKKILIFSKKGYEIG 735

## RESULT 10

US-09-848-909A-13  
; Sequence 13, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0

## RESULT 11

PCT-US03-35733-10  
; Sequence 10, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10

; SEQ ID NO 13  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-13

Query Match 99.4%; Score 3626; DB 23; Length 735;  
Best Local Similarity 96.7%; Pred. No. 4.4e-285;  
Matches 711; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 60

QY 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120

QY 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 180  
DB 121 QRENTEKGLDPLKYWTDSQNKKEVIFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENOYF 180

QY 181 DOIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
DB 181 DOIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420

QY 397 LNAQKASSTPTIWNVQNFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 456  
DB 421 LNAQKASSTPTIWNVQNFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480

QY 457 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 516  
DB 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAGFNEPENG 540

QY 517 QYQKDIETFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAMNILLIRDRPHYDR 576  
DB 541 QYQKDIETFDNFDDQSTQNKQLAELNATNIYVLDKIKLNKAMNILLIRDRPHYDR 600

QY 577 NNIAVGADESVMKEAREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 636  
DB 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660

QY 637 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 696  
DB 661 DMLNSSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 720

QY 697 IKKILIFSKKGYEIG 711  
DB 721 IKKILIFSKKGYEIG 735

;; PRIOR APPLICATION NUMBER: US 60/424,987  
;; PRIOR FILING DATE: 2002-11-08  
;; NUMBER OF SEQ ID NOS: 38  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 10  
;; LENGTH: 735  
;; TYPE: PRT  
;; ORGANISM: Bacillus anthracis  
PCT-US03-35733-10

Query Match 99.2%; Score 3620; DB 1; Length 735;  
Best Local Similarity 96.6%; Pred. No. 1.4e-284;  
Matches 710; Conservative 0; Mismatches 1; Indels 24; Gaps 1;  
  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60  
  
QY 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN 180  
  
QY 181 DGIPLDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
DB 181 DGIPLDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
  
QY 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
DB 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
  
QY 337 NANIRYVNTGTAPIYINVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYINVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
  
QY 397 LNAQKDASTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV 456  
DB 421 LNAQKDFSTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV 480  
  
QY 457 LPOIQETTARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 516  
DB 481 LPOIQETTARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540  
  
QY 517 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPFHYDR 576  
DB 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPFHYDR 600  
  
QY 577 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 636  
DB 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660  
  
QY 637 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 696  
DB 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 720  
  
QY 697 IKKILIFSKKGYEIG 711  
DB 721 IKKILIFSKKGYEIG 735

RESULT 12

US-09-848-909a-10

;; Sequence 10, Application US/09848909a  
;; GENERAL INFORMATION:  
;; APPLICANT: Collier, R. John  
;; APPLICANT: Sellman, Brett R.

;; TITLE OF INVENTION: Compounds and Methods for the Treatment  
;; OF INFECTION: and Prevention of Bacterial Infection  
;; FILE REFERENCE: 00742/060002  
;; CURRENT APPLICATION NUMBER: US/09/848,909A  
;; CURRENT FILING DATE: 2001-05-04  
;; PRIOR APPLICATION NUMBER: US 60/201,800  
;; PRIOR FILING DATE: 2000-09-04  
;; NUMBER OF SEQ ID NOS: 35  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 10  
;; LENGTH: 735  
;; TYPE: PRT  
;; ORGANISM: Bacillus anthracis  
US-09-848-909a-10

Query Match 99.2%; Score 3620; DB 23; Length 735;  
Best Local Similarity 96.6%; Pred. No. 1.4e-284;  
Matches 710; Conservative 0; Mismatches 1; Indels 24; Gaps 1;  
  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60  
  
QY 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPTVPDRDN 180  
  
QY 181 DGIPLDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
DB 181 DGIPLDSLEVEGYTDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDPEKVT 240  
  
QY 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
DB 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
  
QY 337 NANIRYVNTGTAPIYINVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 396  
DB 361 NANIRYVNTGTAPIYINVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
  
QY 397 LNAQKDASTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV 456  
DB 421 LNAQKDFSTPTIMNYNQFLEKTKQLRLDQVYGNIAATYNFENGVRVDTGNSWSEV 480  
  
QY 457 LPOIQETTARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 516  
DB 481 LPOIQETTARIIFNGKDLNVERRIAAYNPSPDLETTKPDMTLKEALKIAFGNEPNGNL 540  
  
QY 517 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPFHYDR 576  
DB 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPFHYDR 600  
  
QY 577 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 636  
DB 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660  
  
QY 637 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 696  
DB 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 720  
  
QY 697 IKKILIFSKKGYEIG 711  
DB 721 IKKILIFSKKGYEIG 735

RESULT 13

PCT-US03-19786-4  
; Sequence 4, Application PC/TUS0319786  
; GENERAL INFORMATION:  
; APPLICANT: Human Genome Sciences, Inc.  
; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof  
; FILE REFERENCE: PV595PCT  
; CURRENT APPLICATION NUMBER: PCT/US03/19786  
; CURRENT FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Mature PA sequence including an ETB signal sequence  
PCT-US03-19786-4

Query Match 99.2%; Score 3617; DB 1; Length 735;  
Best Local Similarity 96.5%; Pred. No. 2.4e-284;  
Matches 709; Conservative 1; Mismatches 1; Indels 24; Gaps 1;  
Db 1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60  
1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKYKKSDEYTFATSDADNHVTMVDQEVINKASNSKIRLEKGRLYQIKIY 120  
61 QSAIWSGFIKYKKSDEYTFATSDADNHVTMVDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKYKKSDEYTFATSDADNHVTMVDQEVINKASNSKIRLEKGRLYQIKIY 120  
121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180  
121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180  
181 DGIPDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
181 DGIPDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETTISKNTSTSRTH 300  
241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETTISKNTSTSRTH 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETTISKNTSTSRTH 300  
301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
337 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 396  
361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 420  
397 LNAQKDSASTPTIMYNOQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 456  
421 LNAQKDSASTPTIMYNOQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
Db 421 LNAQKDSASTPTIMYNOQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
457 LPQIQTETARIIIFNGKDLNVERIAAANVPSDPLETTKPDMTLKEALKIAPGNEPENG 516  
481 LPQIQTETARIIIFNGKDLNVERIAAANVPSDPLETTKPDMTLKEALKIAPGNEPENG 540  
Db 481 LPQIQTETARIIIFNGKDLNVERIAAANVPSDPLETTKPDMTLKEALKIAPGNEPENG 540  
517 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKXNIIIRDKRPHYDR 576  
541 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKXNIIIRDKRPHYDR 600  
Db 541 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKXNIIIRDKRPHYDR 600  
577 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIVIEIDTEGLKEVINDRY 636  
601 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
637 DMLNITSSLRQDGTFFIDPKYNDKPLIYISNPNKYVAVYVTKNTIINPSENGDTSTNG 696  
661 DMLNITSSLRQDGTFFIDPKYNDKPLIYISNPNKYVAVYVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNITSSLRQDGTFFIDPKYNDKPLIYISNPNKYVAVYVTKNTIINPSENGDTSTNG 720  
697 IKKILIFSKKGYEIG 711  
721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

RESULT 14  
PCT-US03-35733-2  
; Sequence 2, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-2  
Query Match 99.1%; Score 3614; DB 1; Length 735;  
Best Local Similarity 96.5%; Pred. No. 4.2e-284;  
Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;  
QY 1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSOGLLYGYFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKYKKSDEYTFATSDADNHVTMVDQEVINKASNSKIRLEKGRLYQIKIY 120  
61 QSAIWSGFIKYKKSDEYTFATSDADNHVTMVDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKYKKSDEYTFATSDADNHVTMVDQEVINKASNSKIRLEKGRLYQIKIY 120  
121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180  
121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKKRSTAGPTVPDRDN 180  
181 DGIPDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
181 DGIPDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETTISKNTSTSRTH 300  
241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETTISKNTSTSRTH 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEQSTQNTDSETTISKNTSTSRTH 300  
301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
337 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 396  
361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQLSOLAPNNYPSKNLAPIA 420  
397 LNAQKDSASTPTIMYNOQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 456  
421 LNAQKDSASTPTIMYNOQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
Db 421 LNAQKDSASTPTIMYNOQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
457 LPQIQTETARIIIFNGKDLNVERIAAANVPSDPLETTKPDMTLKEALKIAPGNEPENG 516  
481 LPQIQTETARIIIFNGKDLNVERIAAANVPSDPLETTKPDMTLKEALKIAPGNEPENG 540  
Db 481 LPQIQTETARIIIFNGKDLNVERIAAANVPSDPLETTKPDMTLKEALKIAPGNEPENG 540  
517 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKXNIIIRDKRPHYDR 576  
541 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKXNIIIRDKRPHYDR 600  
Db 541 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNKXNIIIRDKRPHYDR 600  
577 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIVIEIDTEGLKEVINDRY 636  
601 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSSTEGLLLNDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
637 DMLNITSSLRQDGTFFIDPKYNDKPLIYISNPNKYVAVYVTKNTIINPSENGDTSTNG 696  
661 DMLNITSSLRQDGTFFIDPKYNDKPLIYISNPNKYVAVYVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNITSSLRQDGTFFIDPKYNDKPLIYISNPNKYVAVYVTKNTIINPSENGDTSTNG 720

Db 601 NNTAVGADESIVKAEHREVINSTEGLLLNIDKIDIRKILSGYIVEITEDTEGLKEVINDRY 660  
Qy 637 DMLNSSLRODQGTFFIDFKYNDKLPYISNPNYKNNYAVTKENTIIINPSENGDTSTNG 696  
Db 661 DMLNSSLRODQGTFFIDFKYNDKLPYISNPNYKNNYAVTKENTIIINPSENGDTSTNG 720  
Qy 697 IKKILIFSCKGYEIG 711  
Db 721 IKKILIFSCKGYEIG 735

Search completed: May 3, 2004, 19:57:47  
Job time : 172.814 secs

RESULT 15

US-09-848-909A-2  
; Sequence 2, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-2

Query Match 99.1%; Score 3614; DB 23; Length 735;  
Best Local Similarity 96.5%; Pred No. 4.2e-284;  
Matches 709; Conservative 0; Mismatches 2; Indels 24; Gaps 1;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSPSSSELENIPSENOYF 60  
Qy 61 QSAIWGGFTKVKKSDRYTATSADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120  
Db 61 QSAIWGGFTKVKKSDRYTATSADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKEVISSDNQLQPELKQSSNSRKRKSTAGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKEVISSDNQLQPELKQSSNSRKRKSTAGPTVPDRDN 180  
Qy 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTADPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTADPYSDPEKVT 240  
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Qy 301 -----SNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
Db 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 337 NANIRYVNTGTAPYVNLPTSLVLGKNQTLATIKADENQLSOILAPNNYPSKNLAPTA 396  
Db 361 NANIRYVNTGTAPYVNLPTSLVLGKNQTLATIKADENQLSOILAPNNYPSKNLAPTA 420  
Qy 397 LNAQKASSTPIITMNNYQFLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGNSWSEV 456  
Db 421 LNAQDDPSSTPIITMNNYQFLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGNSWSEV 480  
Qy 457 LPQIQETTARIIFNGKDLNVERRIAAVNPSDPLETTKPDMTLKEALAFAGNEENGNL 516  
Db 481 LPQIQETTARIIFNGKDLNVERRIAAVNPSDPLETTKPDMTLKEALAFAGNEENGNL 540  
Qy 517 QYQGKDIITEFDNFDQTSQNIKNQLAELNATNIYVLDXIKLNAKNILIRKRFHYDR 576  
Db 541 QYQGKDIITEFDNFDQTSQNIKNQLAELNATNIYVLDXIKLNAKNILIRKRFHYDR 600  
Qy 577 NNTAVGADESIVKAEHREVINSTEGLLLNIDKIDIRKILSGYIVEITEDTEGLKEVINDRY 636

hypothetical prote	30	171	4.7	1553	2	T18502
hypothetical prote	31	168.5	4.6	786	2	T18469
probable peptidogl	32	168	4.6	1612	2	AB1347
hypothetical prote	33	166.5	4.6	2526	2	T20531
probable nuclear p	34	166	4.6	1337	2	T41023
hypothetical prote	35	166	4.6	5005	2	F82884
tyb protein - yeas	36	165.5	4.5	1802	2	S52611
probable membrane	37	165	4.5	1308	2	E71622
probable peptidogl	38	165	4.5	2013	2	AD1129
DNA-directed RNA p	39	164	4.5	2910	2	T28156
ORF MSV156 hypothe	40	163	4.5	1127	2	T28317
tyb protein - yeas	41	163	4.5	1465	2	S31262
major merozoite su	42	163	4.5	1831	1	SAZQK1
hemolysin [importe	43	163	4.5	1835	2	AI0452
tyb protein - yeas	44	163	4.5	1803	2	S56894
RhoA-binding prote	45	162.5	4.5	1024	2	T30868

ALIGNMENTS

RESULT 1

I39934

protective antigen precursor - Bacillus anthracis plasmid

C:Species: Bacillus anthracis

C:Date: 19-Jul-1996 #sequence revision 19-Jul-1996 #text\_change 01-Dec-2000

C:Accession: I39934; S69160; F59104

R:Wolkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.

Gene 69, 287-300, 1988

A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr.

A:Reference number: I39933; MUID:89172073; PMID:3148491

A:Accession: I39934

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-764 <RES>

A:Cross-references: GB:M25589; NID:9143280; PIDN:AAA22637.1; PID:9143282

R:Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klumper, K.R.; Birch, N.P.; Loh, Y.P.

Arch. Biochem. Biophys. 316, 5-13, 1995

A:Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SP

A:Reference number: S69160; MUID:95142670; PMID:7840657

A:Accession: S69160

A:Molecule type: Protein

A:Residues: 197-202 <FRI>

R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle

J. Bacteriol. 181, 6509-6515, 1999

A:Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbor

A:Reference number: A59091; MUID:99445483; PMID:10515943

A:Accession: F59104

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-313, 'Q', 315-764 <OKI>

A:Cross-references: GB:AF065404; NID:G4894216; PIDN:AAD32414.1; PID:G4894326

A:Experimental source: strain Sterne

A:Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid

C:Genetics:

A:Gene: pXO1-110

A:Genome: plasmid

C:Function:

A:Description: three component exotoxin; protective antigen binds to receptors on the s

Y active components edema factor or lethal factor; the complex is internalized by recep

C:Keywords: exotoxin

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-156/Domain: propeptide #status predicted <PRO>

F:197-202/Product: protective antigen #status experimental <MAT>

Query Match 98.9%; Score 3607; DB 2; Length 764;

Best Local Similarity 96.3%; Pred. No. 5.2e-173;

Matches 708; Conservative 0; Mismatches 13; Indels 24; Gaps 1;

Qy 1 EVQENRLNESSSQGLLGYFFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 60

Db 30 EVQENRLNESSSQGLLGYFFSDLNFOAPMVVTSITGDLSPSSSELENIPSENQYF 89

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 12.6582 Seconds

(without alignments)

5403.004 Million cell updates/sec

Title: US-09-848-909a-16

Perfect score: 3648

Sequence: 1 EVQENRLNESSSQGLL.....TSTNGIKILIFSKGYEIG 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3607	98.9	764	I39934	protective antigen
2	849	23.3	875	I40862	iota toxin compone
3	235.5	6.5	192	I39933	cryptic protein -
4	235.5	6.5	204	G59104	hypothetical prote
5	200	5.5	4688	JF2885	hypothetical prote
6	197	5.4	1302	1 JC6009	surface-located me
7	192.5	5.3	2529	2 B64635	toxin-like outer m
8	190.5	5.2	2401	2 T28676	rhostry protein -
9	190	5.2	752	2 G90599	hypothetical prote
10	188	5.2	1658	2 S55101	hypothetical prote
11	187	5.1	2269	2 T28677	rhostry protein -
12	187	5.1	6713	2 B89921	hypothetical prote
13	181	5.0	4919	2 T31105	hypothetical prote
14	180	4.9	1227	2 C97033	uncharacterized pr
15	179.5	4.9	624	2 PC5003	surface membrane p
16	179.5	4.9	1365	2 C90822	lmp1 protein - Myc
17	178	4.9	1272	2 C90593	hypothetical prote
18	177.5	4.9	1072	2 A88827	hypothetical prote
19	177	4.9	1839	2 T18372	repeat organellar
20	175	4.8	3216	2 C90538	hypothetical prote
21	175	4.8	3724	2 T18427	hypothetical prote
22	174.5	4.8	821	2 S67087	hypothetical prote
23	174.5	4.8	1115	2 T41342	probable coiled-co
24	173.5	4.8	1125	2 E90598	membrane nuclease,
25	173.5	4.8	1639	2 S05603	major merozoite su
26	173.5	4.8	2399	2 H73879	toxin-like outer m
27	173	4.7	769	2 F89970	serine proteinase
28	171.5	4.7	1208	2 T30668	coiled coil protei
29	171	4.7	1315	2 T28679	fibrinogen-binding

QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 90 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 149  
QY 121 QSENTEKGLDFKLYWTDSONKKEVYSSDNQLQLPELKQKSNRKRKSTSGPTVPDRDN 180  
DB 150 QSENTEKGLDFKLYWTDSONKKEVYSSDNQLQLPELKQKSNRKRKSTSGPTVPDRDN 209  
QY 181 DGIPOSLEVEGYTVVQKRTFLSPWNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
DB 210 DGIPOSLEVEGYTVVQKRTFLSPWNIHEKGLTKYKSPKSWSTASDPYDFEKT 269  
QY 241 GRIDKNVSPKAPHPVAAPIVHVDMENIILSKNEDQSTQNTDSTRTISKNTSRTHT 300  
DB 270 GRIDKNVSPKAPHPVAAPIVHVDMENIILSKNEDQSTQNTDSTRTISKNTSRTHT 329  
QY 301 -----SNSNSTVAIDHSLSLAGERTWAGTMTGLTADTARL 336  
DB 330 SEVHGNAEVHAFDIDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAGTMTGLTADTARL 389  
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYYPKSNLAPIA 396  
DB 390 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYYPKSNLAPIA 449  
QY 397 LNAQKQDASPTTMYNQVQLEKTKQLELDQVYGNATYVNGRVRVDTGSKNSEV 456  
DB 450 LNAQKQDASPTTMYNQVQLEKTKQLELDQVYGNATYVNGRVRVDTGSKNSEV 509  
QY 457 LPQIQTETARIIFNGKDLNVLERRIAAANVPSDPLETTKPDMLKALKAIFGNEPENG 516  
DB 510 LPQIQTETARIIFNGKDLNVLERRIAAANVPSDPLETTKPDMLKALKAIFGNEPENG 569  
QY 517 QYQKQDITFDNFDOOTQNTKNQALBELNATYVLDKIKNAKMMILIRDKRPHYDR 576  
DB 570 QYQKQDITFDNFDOOTQNTKNQALBELNATYVLDKIKNAKMMILIRDKRPHYDR 629  
QY 577 NNIAVGADESVMKEAAREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 636  
DB 630 NNIAVGADESVMKEAAREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 689  
QY 637 DMLNITSLQDQKTFIDFKYNDKPLXYISNPNYKNNVATYKNTIINPSENGDTSTNG 696  
DB 690 DMLNITSLQDQKTFIDFKYNDKPLXYISNPNYKNNVATYKNTIINPSENGDTSTNG 749  
QY 697 IKKILFSSKGYEIG 711  
DB 750 IKKILFSSKGYEIG 764

RESULT 2  
I40862  
iota toxin component Ib - Clostridium perfringens  
C:Species: Clostridium perfringens  
C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Oct-1999  
C:Accession: I40862; S42774  
R:Perelle, S.; Gibert, M.; Boquet, P.; Popoff, M.R.  
Infect. Immun. 61, 5147-5156, 1993  
A:Title: Characterization of Clostridium perfringens iota-toxin genes and expression in  
A:Reference number: I40861; MUID:94041637; PMID:8225592  
A:Accession: I40862  
A>Status: preliminary; translated from GB/EMBL/DDSI  
A:Molecule type: DNA  
A:Residues: 1-875 <RES>  
A:Cross-references: EMBL:X73562; NID:g929031; PID:CAA51960.1; PID:g414655

Query Match 23.3%; Score 849; DB 2; Length 875;  
Best Local Similarity 31.3%; Pred. No. 5, 8e-35;  
Matches 251; Conservative 128; Mismatches 268; Indels 154; Gaps 33;

QY 1 EVQENRLLNESSSQGLLGYYFSDINFOAPMVVTSSTGDLISIPSSLENIPE-NQY 59  
DB 35 DTNQEKEITTENTILSSNGLMGYYFADEHFKDLKPEKVKVLLTEDNSS 94

QY 60 FOSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 119  
DB 95 IKSIRWTGRIIPSEDEGEYILSTR-NDVLQINAKGDIKAK---TLKVNKKQKQAYNIRIE 150  
QY 120 YQREN-----PTEKGLDFKLYWTDSONKKEVYSSDNQLQLPELKQKSNRKRKST 170  
DB 151 IQDKNIGSIDNLSVP-----KLYW-ELUNGKRTVPEENLFRDYSKIDEND----- 195  
QY 171 AGPTVP-----DRNDGIFPDSLEVEGYTVVQKRTFLSPWNIHEKGLTKYKSPKSW 211  
DB 196 --PFINNFFDVRPFSAAWEDDLTDNDNIPDAVEKNGYTI---KDSIAVKNDSPA 249  
QY 212 EKGLTKYKSPKSWSTASDPYDFEKTGRDKQVSPKAPHPVAAPIVHVDMENIIL 271  
DB 250 E-QGYKYYSSVLESNTAGDPYDYQKASGSDKAIKLEARDPLVAAVYVGVGMENTLI 308  
QY 272 SKVE-----DOSTQNTDSETRTI-----SKNTSRTHTSNTSSTVA 309  
DB 309 STNEHASSDQKTVSRAITNSKTDANTVGVSIAGYQNGFTGNTIITSYHT--TDNSTAV 366  
QY 310 IDHLSLAGERTWAGTMTGLTADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKQTLAT 369  
DB 367 QDSN---GE-SMNTGLSINKGESAYINANVRYNTGTAPMYKVTPTTNLVL-DGETLAT 420  
QY 370 IKADENQLSQILAPNNYYPKSNLAPIALNAQKQDASPTTMYNQVQLEKTKQLELD 429  
DB 421 IKADENQLSQILAPNNYYPKSNLAPIALNAQKQDASPTTMYNQVQLEKTKQLELD 480  
QY 430 QYVGNATYVNGRVRVDTGSKNSEVLPQIQTETARIIFNGKDLNVLERRIAAANVPSD 489  
DB 481 QVSGNYGTYK-SQGI-ITEGNSWNYISQIDSVSASIIID-TGSOTFERRVAAKEQGNP 537  
QY 490 LETTKPDMLKALKAIFGNEPENGQY-QGKQITE--PDFNFDOOTQNTKNQALBELN 546  
DB 538 EDKT-PEITTEGEAIIKAFSAATK-NGELLYFNGIPIDESCVELIPDNTSEIKEQLKYLD 595  
QY 547 ATNIYTVLDKIKNAKMMILIRDKRPHYDR-NNIAVGADESVMKEAAREVINSSTEG 603  
DB 596 DKKIYV---KLERGMNIIKVPYSYFTNPFDEYNNFP--ASWSNIDTKNQDGLQSVANKL 649  
QY 604 -----LLNDKDKIRKILSGY-----IVEIEDTEGLKEVINDRYDMLNIS-- 642  
DB 650 SGETKIIIPMSKULPKRYVFGYSGYKDPSTNSITVNIKSKEQKTYLVPKDYTFESYE 709  
QY 643 -----SLRQDQKTFIDFKKYND--KLPLYISNPNYK---NVYA 676  
DB 710 FETTKDSSDIETLTSSGVIPLDNLISITELNTPBILKEPKVPSDQIBLDHAKNYA 769  
QY 677 VTENTIIINPSENGDTSTNGI 697  
DB 770 DIKLDI---NTGNTYIDGI 785

RESULT 3  
I39933  
cryptic protein - Bacillus anthracis  
C:Species: Bacillus anthracis  
C>Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 15-Oct-1999  
C:Accession: I39933  
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.  
Gene 69, 287-300, 1988  
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr  
A:Reference number: I39933; MUID:89172073; PMID:3148491  
A:Accession: I39933  
A>Status: preliminary; translated from GB/EMBL/DDSI  
A:Molecule type: DNA  
A:Residues: 1-192 <RES>  
A:Cross-references: GB:M22569; NID:g143280; PID:AAA22636.1; PID:g143281

Query Match 6.5%; Score 235.5; DB 2; Length 192;  
Best Local Similarity 34.6%; Pred. No. 3e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 563 MNILIRDKRPHYDRNNAVGADESUVKAEHREVINSSTEGLLNIDKIRKILSGYVEI 622  
DB 1 MNILVRDP-YHYDNNNGNIVGVDDSYLKNAYKQILNWSGVSGLNLDVQALSGYMLQI 59  
QY 623 EDTE-----GLKEVINDRYDMLNMISSLRQDGKTFIDFKYNDKXLPYISNPN 669  
DB 60 KPSNHLTNSPVTITLACKDSGVGELYRVLS-----DGAGFLDFNKFEDNWRSLV-DPG 112  
QY 670 YKXNVYAVTKEN-TIINPSENGDTSWGKIKLIFSKGYEI 710  
DB 113 DDVYVYAVTKEDFNAVTRDENGNTA-NKLKNTLVLSGKIKEI 153

RESULT 4  
G59104  
hypochemical protein pX01-111 - Bacillus anthracis virulence plasmid pX01  
C:Species: Bacillus anthracis  
C>Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 11-May-2000  
C:Accession: G59104  
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hall, K.K.; Keim, P.; Koehler  
J. Bacteriol. 181, 6509-6515, 1999  
A:Title: Sequence and organization of pX01, the large Bacillus anthracis plasmid harbored  
A:Reference number: A59031; MUID:99445483; PMID:10515943  
A:Accession: G59104  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-204 <OKI>  
A:Cross-references: GB:AF065404; NID:g4894216; PIDN:AAD32415.1; PID:g4894327  
A:Experimental source: strain Sterne  
A:Note: Similar to hypothetical protein in the protective antigen domain; ypa, plasmid p  
C:Genetics:  
A:Gene: pX01-111  
A:Genome: plasmid

Query Match 6.5%; Score 235.5; DB 2; Length 204;  
Best Local Similarity 34.6%; Pred. No. 3.3e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 563 MNILIRDKRPHYDRNNAVGADESUVKAEHREVINSSTEGLLNIDKIRKILSGYVEI 622  
DB 1 MNILVRDP-YHYDNNNGNIVGVDDSYLKNAYKQILNWSGVSGLNLDVQALSGYMLQI 59  
QY 623 EDTE-----GLKEVINDRYDMLNMISSLRQDGKTFIDFKYNDKXLPYISNPN 669  
DB 60 KPSNHLTNSPVTITLACKDSGVGELYRVLS-----DGTGFLDFNKFEDNWRSLV-DPG 112  
QY 670 YKXNVYAVTKEN-TIINPSENGDTSWGKIKLIFSKGYEI 710  
DB 113 DDVYVYAVTKEDFNAVTRDENGNTA-NKLKNTLVLSGKIKEI 153

RESULT 5  
F82885  
hypothetical protein UV482 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: F82885  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, B.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
A:Reference number: A82870  
A:Accession: F82885  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-468 <GLA>  
A:Cross-references: GB:AE002145; GB:AF222894; NID:g6899476; PIDN:AAF30894.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UV482  
A:Genetic code: SGC3

Query Match 5.5%; Score 200; DB 2; Length 468;  
Best Local Similarity 20.3%; Pred. No. 0.2;

Matches 166; Conservative 127; Mismatches 313; Indels 210; Gaps 37;  
QY 19 LLGYFSDLNFGAPMVVTSSTTGDLSIPSSLEENIPSENQVFSQAIWSGFIKVKKSDYET 78  
DB 3699 LVDDVYLD-NHONIDEITRKFQKDNV-SKIEINPGVTMLSKHGNWKSPTDTTANFEFK 3756  
QY 79 FATSADNHVTWMDVQEVINKASNKRLKGRUYQIKIQORENPTEKGLDFKLYWTD 138  
DB 3757 IETQ-----DDNDVLNNDATVFKPKDEHNNTKQKIVRIKEN-----ND 3795  
QY 139 SQNKKEVISSDNLQLPKQKSSN---SRKGRSTSAGTVPDRD----- 179  
DB 3796 WLIKQI---DNLN-PETKYKLENTLSKPLKTHNLSVINDKENISLITETGNPVLKV 3851  
QY 180 ----NDGIPDSLEVEGYTDVKNKRTFLSPWTSNIHEKKGLTKYKSSPEKWSGADSPVD 235  
DB 3852 IQQNDTINDTQTTINVTLSGVNSK-YNGRQKQVYKNNNNVIYESS---LITLQKQND 3907  
QY 236 FEKVTGRIDKQVSPPEARHPLVAAPYIVHVDMENIILSKN-EDQSTQNTDSEY-RTISKNT 293  
DB 3908 YQLLSNLNSN-----REYAFKEIENHISNTNNFEDLEKLVGSNFTITQTKNT 3957  
QY 294 STSRHTSNSSSTVAIDHSLSLAG----- 318  
DB 3958 TVQWDDSSATVGTGCVNFNFKISBKDKILENNQVAVFAPKETIRDTNTWLQYTRPLK 4017  
QY 319 -----ERTWAEITMGLNT---ADTAPLNANTRYNTGTAPIYVNLPTTSLVLGKQNTLA 368  
DB 4018 DVTSDFKETGWAHLDSNVFKEETTYKLVIQFVKNKPTAKNNNNNSNNVILONTNSI 4077  
QY 369 -----TIKADENQLSCILAPNNYPSKNLAPIALN-AQKDASSSTPITWNY--NOFLELE 419  
DB 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTIINFTLSGVKKSVMGKKIKLSYKSNDSSES 4137  
QY 420 KTKQLRLDQDVVGNITATNFENG-----VRVDTGSNWSEVLQIQIETTARILFNGKD 473  
DB 4138 HTNEVLIESNKTQYNTLLNNLKNRTYTLIDVKLIDNNNVSDFPKKEGNTJNSFITTRTSA 4197  
QY 474 LNLVERRIAAVNSDPLETT-----KPDMLTKEA-----LKIAGFNEPENGNIQ 517  
DB 4198 INVITEEISNRASTNLKSTIINKINLNDPDLVLDKQATIVYGNKNGAMGFTIVSGNIK 4257  
QY 518 YQKGDITEFDNPDQOTSONIK-NQAEALNATNYIVLDKILKNAKONLILDRKPHYD- 575  
DB 4258 YLTATLVDLNFN-DKVNIVNISFNKPSAAEN-----IGDKSNII-----YNNDS 4304  
QY 576 -----RNNTAVGA---DESUVKAEHREVINSSTEGLLNIDKIRKILSGYVEI 626  
DB 4305 IPKLEINNDIIVNGPINKELIVKRAQK--NNIDVDLGLQINPKIAHLR-PIAKFKSTN 4361  
QY 627 GLKEVINDD--RYDMLNISSL-RQDGKTFIDF-----KKY----- 657  
DB 4362 -----NDIETVINGSSLVNNDGKTSIRFTLNLKANKLYSLVDVYVIVNNSNTIVE 4415  
QY 658 NDKLPLYISNPNYKXNVYAVTKENTINPSENGDTS 693  
DB 4416 SNKLP-KLANNINYQIK---INKSHTII--SKNGEWS 4445

RESULT 6  
JC6009  
surface-located membrane protein lmp3 precursor - Mycoplasma hominis  
C:Species: Mycoplasma hominis  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: JC6009  
R:Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkegaard, S.; Christiansen, C.  
J. Bacteriol. 178, 2775-2784, 1996  
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene syste  
A:Reference number: JC6009; MUID:96213016; PMID:8631664  
A:Accession: JC6009  
A:Molecule type: DNA  
A:Residues: 1-1302 <LAD>  
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336

Matches 166; Conservative 127; Mismatches 313; Indels 210; Gaps 37;  
QY 19 LLGYFSDLNFGAPMVVTSSTTGDLSIPSSLEENIPSENQVFSQAIWSGFIKVKKSDYET 78  
DB 3699 LVDDVYLD-NHONIDEITRKFQKDNV-SKIEINPGVTMLSKHGNWKSPTDTTANFEFK 3756  
QY 79 FATSADNHVTWMDVQEVINKASNKRLKGRUYQIKIQORENPTEKGLDFKLYWTD 138  
DB 3757 IETQ-----DDNDVLNNDATVFKPKDEHNNTKQKIVRIKEN-----ND 3795  
QY 139 SQNKKEVISSDNLQLPKQKSSN---SRKGRSTSAGTVPDRD----- 179  
DB 3796 WLIKQI---DNLN-PETKYKLENTLSKPLKTHNLSVINDKENISLITETGNPVLKV 3851  
QY 180 ----NDGIPDSLEVEGYTDVKNKRTFLSPWTSNIHEKKGLTKYKSSPEKWSGADSPVD 235  
DB 3852 IQQNDTINDTQTTINVTLSGVNSK-YNGRQKQVYKNNNNVIYESS---LITLQKQND 3907  
QY 236 FEKVTGRIDKQVSPPEARHPLVAAPYIVHVDMENIILSKN-EDQSTQNTDSEY-RTISKNT 293  
DB 3908 YQLLSNLNSN-----REYAFKEIENHISNTNNFEDLEKLVGSNFTITQTKNT 3957  
QY 294 STSRHTSNSSSTVAIDHSLSLAG----- 318  
DB 3958 TVQWDDSSATVGTGCVNFNFKISBKDKILENNQVAVFAPKETIRDTNTWLQYTRPLK 4017  
QY 319 -----ERTWAEITMGLNT---ADTAPLNANTRYNTGTAPIYVNLPTTSLVLGKQNTLA 368  
DB 4018 DVTSDFKETGWAHLDSNVFKEETTYKLVIQFVKNKPTAKNNNNNSNNVILONTNSI 4077  
QY 369 -----TIKADENQLSCILAPNNYPSKNLAPIALN-AQKDASSSTPITWNY--NOFLELE 419  
DB 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTIINFTLSGVKKSVMGKKIKLSYKSNDSSES 4137  
QY 420 KTKQLRLDQDVVGNITATNFENG-----VRVDTGSNWSEVLQIQIETTARILFNGKD 473  
DB 4138 HTNEVLIESNKTQYNTLLNNLKNRTYTLIDVKLIDNNNVSDFPKKEGNTJNSFITTRTSA 4197  
QY 474 LNLVERRIAAVNSDPLETT-----KPDMLTKEA-----LKIAGFNEPENGNIQ 517  
DB 4198 INVITEEISNRASTNLKSTIINKINLNDPDLVLDKQATIVYGNKNGAMGFTIVSGNIK 4257  
QY 518 YQKGDITEFDNPDQOTSONIK-NQAEALNATNYIVLDKILKNAKONLILDRKPHYD- 575  
DB 4258 YLTATLVDLNFN-DKVNIVNISFNKPSAAEN-----IGDKSNII-----YNNDS 4304  
QY 576 -----RNNTAVGA---DESUVKAEHREVINSSTEGLLNIDKIRKILSGYVEI 626  
DB 4305 IPKLEINNDIIVNGPINKELIVKRAQK--NNIDVDLGLQINPKIAHLR-PIAKFKSTN 4361  
QY 627 GLKEVINDD--RYDMLNISSL-RQDGKTFIDF-----KKY----- 657  
DB 4362 -----NDIETVINGSSLVNNDGKTSIRFTLNLKANKLYSLVDVYVIVNNSNTIVE 4415  
QY 658 NDKLPLYISNPNYKXNVYAVTKENTINPSENGDTS 693  
DB 4416 SNKLP-KLANNINYQIK---INKSHTII--SKNGEWS 4445





T28676  
rhoptry protein - Plasmodium yoelii (fragment)  
C:Species: Plasmodium yoelii  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T28676; A45521  
R:Singha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.  
Mol. Biochem. Parasitol. 76, 329-332, 1996  
A:Title: Comparison of two members of a multigene family coding for high-molecular mass  
A:Reference number: 220507; MUID:97077455; PMID:8920022  
A:Accession: T28676  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2401 <SIN>  
A:CROSS-references: EMBL:U36927; NID:g1041784; PID:g1041785; PIDN:AAB41263.1  
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-245, 1990  
A:Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co  
A:Reference number: A45521; MUID:91101860; PMID:2270106  
A:Accession: A45521  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 2260-2401 <KEP>  
A:CROSS-references: GB:M34281

Query Match 5.2%; Score 190.5; DB 2; Length 2401;  
Best Local Similarity 21.6%; Pred. No. 0.22;  
Matches 176; Conservative 120; Mismatches 321; Indels 197; Gaps 40;

QY 1 EVKQENLLNESSSSQGLGYFSDLN-----FCAPVVTSSITGDLIS 45  
DB 402 EVAKENVQNVYKSNILEIKKYNQINIDNKEAKQNDQFKHKMTTPNEMKYQK 461

QY 46 PSELENIPSENYFQSAL--WSGFIK---KKSDEYTFATSADNHVTWVDDQEV---I 97  
DB 462 PSIEIKIMKDE--FLSKVKNYNDPKVYKVESEHKNFTLTNKTIVTSDEBIKKYE 518

QY 98 NKASNSNKIRLEKGRLYKIQYQRENPEKGLDKFYWDSONKKEVISSDNQLPLK 157  
DB 519 NKFDNSKSLNETK---STEEYQNTLTKVDD---YIKVCLNTNELITNCHNQITLK 573

QY 158 QKSSNSRK--KRSTAGTPVPDRDNDGIPD-----SLEVEGYVDVKNKTFPLSP 205  
DB 574 DKLNQNIKITEKNSIDKIYTDKFNELITDKKTELEKTFGLSLNHNESNNKELLTYFD 633

QY 206 WTSNHEKGLTYKSSPKNSASDPYDFKVTGRDKNVSPARHPLVAAYPIVHVD 265  
DB 634 LXANLGKKNENLYKOFNEK---EKAVEDIRKKNVNDINKIVSN-----IRIT 677

QY 266 MENILSKNEDOSTNDSETRTISKTSTSRTH7SNSNSTVAI--DRSLSLAGERTWAE 324  
DB 678 IYTSIYNINEDTENEIGKSIELLNTKYLEKVKANVTNLNEIKELKDYDFQDFGK---- 732

QY 325 TWGLNTADTARLANANRYNTGTAPIYVLPITSLVLGN--OTLATIKAD--ENQUSQI-- 380  
DB 733 -----EKNIKYPDEN--KIKNDIDTLNQIKDSIETLTTEIKKNSENHIDEIKG 778

QY 381 -----LAPN-----SKNAPIALNAQ-----DASSTPTWNYNQFLE 419  
DB 779 QIDKLKVPKNTWNPEDPKEIEKKIENIVEKDKKNYIKEYDKLLNELSKIENDKTSLE 838

QY 420 KTKQLRLDTDOVYGNATYFNGRVRVDTGSNWSEV-----LPOIQTETARIIFNGKLN 475  
DB 839 KLANINLSYKSLGNLFQIODEKKKAETHIKAMEAYDLDNINKKXSOEI---EKEMN 895

QY 476 L-----VERRTAANPSDP-----LETT-----KPDMLKEALKIAGFNEP---NGN 515  
DB 896 INMDIKDTHKWKALNISHDDYKIVHTTSKNHEEKISDIRKNSLKIQDFSEESYINDI 955

QY 516 LOYQGGDIETFDNFPQOQSNTIKNLAELNATNYTVLDDKIKNAKNNILRDRKRF--H 573  
DB 956 KKELEKNVLE-----SQNNNTDINQYLSKIE--NIYNIL--KLNKIKKIIDKVEYDTE 1005

QY 574 YDRNNIAGADESVVKEAHRVINSSTEGLLNIDKDIRKILSGYVIEIEDTEGLKE--- 630

Db 1006 IEKNN-----KKINAEIENS-----EKI-----ITQKENSLSKECQS 1038  
QY 631 ----VINDRY---DMINISLRODGTDFDKYKNDKPLIYISN--PNYKVN----- 674  
Db 1039 KIKSTIDDNVSECIKNITNL-----KTYIVNEKN--INTYFKABEYNQVSLNFNIE 1092  
QY 675 YAVTKENTINPSENGDTSTN--GIKKILIFSCK 706  
Db 1093 MADTKSQYILNKKNGTNTNDYNIKELKEHKK 1126

RESULT 9  
G90599  
hypothetical protein MYPV 7030 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C:Species: Mycoplasma pulmonis  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: G90599  
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, P.; Moszer, I.  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul  
A:Reference number: A9512; MUID:21267165; PMID:11353084  
A:Accession: G90599  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-752 <KUR>  
A:CROSS-references: GB:AL445566; PID:g14090118; PIDN:CAC13876.1; GSPDB:GN00153  
A:Experimental source: strain UAB CTIP  
C:Genetics:  
A:Gene: MYPV 7030  
A:Genetic code: SGC3

Query Match 5.2%; Score 190; DB 2; Length 752;  
Best Local Similarity 18.7%; Pred. No. 0.043;  
Matches 152; Conservative 141; Mismatches 258; Indels 260; Gaps 41;

QY 41 GDLSISSELENIPSENYFQSALWSGFIKVKKSDEYTF-----A7SADNHVTWVD 92  
DB 20 GTISGYGISQLPNESASL-----VKAHELMFNLNKQVLYNKSIEELINQWQD 69

QY 93 DQEVINKASN-----SNKIRLEKGR--LYQI---KIQYQRENPEKGLDKFYWTDQ 140  
DB 70 ENKVNISANDFFNKVFSQKTPLENGEKITYSVLGRDIFVQIWNPSQKITSKVISSK 129

QY 141 NKEVTSNDLQPLPELKQKSSN--SRKSTSTAGTPVPDRDND---GTPDSLEVEGYTV 194  
DB 130 ISKDVWMD-----KQKELNDFAKNLRVNFKGSASEKQSDIWAQSFNDKSKLBKYL 182

QY 195 DVKN-----KRTFLSPWISNIHEKK---GLTKYKSPKEMSTASDPY 233  
DB 183 DDKNVEKNISENTEFDELKTEENAFVGGYSNDIAGTKVIEAIVKYNSEK-----Y 235

QY 234 SDFEKTGRIDKN-----VSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETR 288  
DB 236 QNIQKI--QITNNFKRFDTSDES-----LRWELSNVKTERTQSKSEN---DASA 280

QY 289 ISKNTSTSRTH7SNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARLANANRYVNTGTA 348  
DB 281 FSGNSLVSKNSLKNDEKIKLESFV-----FDLKTNTKLSVKYADIK----- 327

QY 349 PIYNVLPTTS-----LVLGKNQTLATIKADENQSLQILAPNPNYPSKNLAP 394  
DB 328 --FNKLETVGEQKSVKLTLYLIGWKVDGNESTI-----PNRLTSLFRPS---SKESTI 375

QY 395 IALNAQKASSTPTWNYNQFLEKTKQLRLDTDOVYGNATYFNE-----NGRVRVDTG 450  
DB 376 ILTGLKEELREKLLNSIVDKVELKWTSKDAID--KILASTITSKDFQSQANKRITIGPK 434

QY 451 SN-----WSEV--LPOIQTETTA-----RIIFNGKDLNMLVERRIAAV--NPSPDLETTK 494  
DB 435 ANTDIAKELAKYLTEISVENINDVGTLYLVKLVFLKDDSTINRTITLVLGSGFAKVEATK 494

QY 495 PD-----MTLKEALKIAGFNEPFGNGLQYQGGDIETFDNFPQOQSNTIKNLAELNAT 548



```
Db 702 NKLSEISKEDNTSLEKVKDINLSY--GQNLGNLFLE---QIDEEKKKAENTIK----- 752
Qy 170 SAGTVPDRDN-----DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPE 224
Db 753 SMEAYIDDLNIIKKKSQETETMDIK--MDINKWEALK--ISHDDDKCHKDKSKHKE 807
Qy 225 KWSTASDP-----YSDFEKVTGRIDKNVSPAEARH-----PLVAAPIYHVD 265
Db 808 NISDIYKSKKLIQDPSRESNDINKLQKVNSESQNHSDINOCLENAVANIYNILKN 867
Qy 266 MENIILSKNEDOSTQNTSETRTISKNTSTSRHTSNSSNSTVAIDHLSLAGERTWAET 325
Db 868 KIKKIIDKVKETSE-----IEKNKKNINDELNKGKVIKIEGDLSECKRSKIN- 918
Qy 326 MGLNTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKNGQTLATIKADENQLSQILAPNN 385
Db 919 ---STLDDKDIDECINLN-----VLKKN----- 939
Qy 386 YYPKSLAPIALNAOKDASSTPITMYNOFLEKTKQLRLDQVGYNIATYNPENGRV 445
Db 940 -----ILNEE---TNIT--NHFKNAEB-----YNKIVLSNFFN-- 967
Qy 446 RYDTGSNNSEVLPOIQTETARIIIFNGKDLNLVERRIAIVNPSDPLETTKPDWTLKALKI 505
Db 969 -LEMADNKSQYILEIKONGT---NDHDYNIKELK-SHKDKSNGYK-TEADQNKKAIOK- 1020
Qy 506 AFGFNEPNGNLOVQKGD-----ITEPDPNFDOQTSONIKQLAELNATNIYTVLD 555
Db 1021 -----NKELFEQYKEEVTVLLNKYYAVELKNKFD-KTKDQSKQIIEIKDAHNYCTLE 1072
Qy 556 KIKLNAKON-----ILIRDKRPHYDRNNAIVGADSVVKEAHREVIN-----SSTEGLLL 605
Db 1073 SGKSEKONNEIKNEKIHIEDEVANDKSKATISIKSVPEPKTKIIEINIRTKSDDC 1132
Qy 606 NIDKDKIRKILSGYVIEIDT-----EGLKEVINDRYDMLNLSLRQDGKTFIDFKK 656
Db 1133 KETNLEKQISNLSIDTQETKLTENGKQKLTLEL-----LESKKQKKNIEDQKK 1183
Qy 657 YNDKJPLYISN-----PNYKNVAVTKENTIIINPSNGDTSTNGIKK 699
Db 1184 ELDEVNSKIKNIENTVNOHKNYEIGI--VEKINSIAKTNKQIBSTKELIK 1233

RESULT 12
B89921
hypothetical protein ebhA [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Qui, L.; Oguc
ma, A.; Mizurani-Uji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A>Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89921
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUP>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BA842527.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
A:Gene: ebhA

Query Match 5.1%; Score 187; DB 2; Length 6713;
Best Local Similarity 20.6%; Pred. No. 1.5;
Matches 170; Conservative 127; Mismatches 289; Indels 238; Gaps 41;

Qy 9 LNESESSSQGLLYFSDLNFOAPMVVTS-----STTGDLSPSPSELENIIPSENYFOSA 63
Db 2468 VQAKSDAKANLG-TLTHLNNNAQKQDLTSQIEGATVGVNSVTKAQDLDGAMORLESA 2526
Qy 64 IWSGRIKVKKSDEY-----TFATSADNHVTVMWDDQEVINKASNNKIRLEKRLYQIKI 118
```



N:Alternate names: hypothetical 624 protein; Lmp4 protein  
C:Species: Mycoplasma hominis  
C:Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 11-Jan-2002  
C:Accession: PC6003  
R:Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.  
J. Bacteriol. 178, 2775-2784, 1996  
A:Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene system  
A:Reference number: JG6009; MUID:96213016; PMID:8631664  
A:Accession: PC6003  
A:Molecule type: DNA  
A:Residues: 1-824 <LAD>  
A:Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64859.1; PID:g1197337  
C:Genetics:  
A:Gene: lmp4  
C:Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homology  
C:Keywords: duplication; membrane protein

Query Match 4.9%; Score 179.5; DB 2; Length 624;  
Best Local Similarity 20.8%; Pred. No. 0.11;  
Matches 143; Conservative 109; Mismatches 245; Indels 189; Gaps 34;

QY 81 TSADNHVTMMVDDQEVINKASNSKIRLEKGR-L-YQIKIQYORENPTEKGL-----D 131  
DB 6 SSLDAKVTEITKLETFKDKDKVFKLEQTRKIDIDEFINTKTNPNYSILSELTSKRD 65  
QY 132 FKLWYDSONKKEVISSNLOLPELQ-----KSSNSRKRKSTSGAPT- 174  
DB 66 SKNSITNSKSDIETANT-----ELKQALAKANTDKAQADNLARSTKEQLNKSISSANTL 121  
QY 175 ---VPRDNDGIPDSLEYV-----GYTVDVKNKETFLSPWISNIHEKKGLTKY 219  
DB 122 LAKLTKDONTTQQAQTEKEVQKQAVASNNNTASMSKSSLDKAVTEI-TKKLETFN 180  
QY 220 KSSPEKSTASDPYDFEKVTRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQST 279  
DB 181 KDKDAK-----FKELQTRKNIDEFINTKTNP-----NYSILI-----SOLT 218  
QY 280 QNTDSETRISKNTSTSTHTSNSSSTVAIDHSLSLAG-ERTWAEFMGLNTADTARINA 338  
DB 219 SKRD-----SKNSVTNSKSDIETANTLQALAKANTDKQADNLAKSAKE--QLNN 270  
QY 339 NIRVYNTGTAPIYVLP\*TSVLGKNOTLATIKADENQLSOILAPNNYPPSKNLAPI-AL 397  
DB 271 SVSSANTLLAKLTD-----KDNTIQAKTELEK--EVQKQAVASNNNTASMSQSA 318  
QY 398 NAKDASSTPTM---NYN-----QFLEKTKQLRLDQVYGNIAATYNFENGVRVDT 449  
DB 319 KSSLDKAVTEITKLETFKDKDKVFKLEQTRKIDIDEFINTKTNPNYSILSELTSKRD 65  
QY 450 GSNWSEVLPIQETTARIIFNGKDLNLVERIAAVNPSDPLETTKPDMTLKEALKIAPGF 509  
DB 365 -----PQTKKNYQNVKVLKDKKAEKNSITFSN--NKKEIQDANKSLQDEL----- 408  
QY 510 NEPNGLQVQGDITEFDNPFQQTQSNIKQOLA-ELNATNIYVVDKIKLNKKNILIR 568  
DB 409 ----NNAKITTKGITDF-YNSKKQLEDLITDQDAKVGTTTEADTILD----- 450  
QY 569 DKRFHYDRNNIAGDAESVWKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGL 628  
DB 451 ----HY--KNISDAKNEBEIKQATQKI-----NDIKKIIETKIQE----- 484  
QY 629 KEVINDRYDMLNISLRQDKFTI--DFK--KYNDKPLIYISNPNYKVNYYAVTKENTII 684  
DB 485 ----KKRNEFSQEQIKNELQSFINKDLKQKYN-SIRTKIEN---KINGVSSINKNSKI 536  
QY 685 NPSENG----DTSTNGIKKILIFSCK 706  
DB 537 QDIENAKKOLEFAKNEFEKIINNIKK 562

Search completed: May 3, 2004, 19:41:26  
Job time : 15.6582 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 7.98439 Seconds  
(without alignments)  
4636.784 Million cell updates/sec

Title: US-09-848-909a-16

Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKKILFSSKGYEIG 711

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3604	98.8	764	1	PAG_BACAN
2	235.5	6.5	204	1	YPI1_BACAN
3	198	5.2	1658	1	YME7_YEAST
4	173.5	4.8	1830	1	MSPI_PLAFK
5	173.5	4.8	1639	1	MSPI_PLAFW
6	171.5	4.7	1208	1	PCP1_SCHPO
7	163	4.5	1803	1	YUL3_YEAST
8	162.5	4.5	1024	1	RIP3_MOUSE
9	162.5	4.5	1233	1	YPI6_YEAST
10	161	4.4	2867	1	RB22_PLAVB
11	160.5	4.4	1487	1	MD33_YEAST
12	160.5	4.4	2869	1	RBPI_PLAVB
13	159.5	4.4	1162	1	EXEN_CLOBU
14	159.5	4.4	1358	1	SIR4_YEAST
15	159.5	4.4	2116	1	MY82_DICDI
16	158	4.3	1790	1	US01_YEAST
17	157	4.3	1000	1	S155_YEAST
18	157	4.3	1029	1	RIP3_RAT
19	156	4.3	1957	1	SPOF_SCHPO
20	155.5	4.3	1167	1	CAGA_HELPJ
21	155	4.2	1385	1	FAT1_SCHPO
22	153.5	4.2	1046	1	SBCC_LACLA
23	152.5	4.2	730	1	GLN3_YEAST
24	152.5	4.2	895	1	RASO_THEVO
25	152.5	4.2	2339	1	RPCL_PLAFW
26	151	4.1	2167	1	BEM2_YEAST
27	150.5	4.1	1005	1	RA50_MBTJA
28	150.5	4.1	1570	1	P3X1_DICDI
29	150	4.1	3418	1	BR22_HUMAN
30	149.5	4.1	1162	1	EXEN_CLOBU
31	149.5	4.1	2376	1	TA03_YEAST
32	149	4.1	719	1	YMA1_YEAST
33	149	4.1	976	1	SCPI_HUMAN

#### ALIGNMENTS

```

RESULT 1
PAG_BACAN
ID PAG_BACAN STANDARD; PRT; 764 AA.
AC P13423; Q9F5R7; Q9KH69; Q9ROU2;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins-
DE translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].
GN PAGA OR PAG OR PX01-110.
OS Bacillus anthracis.
OG Plasmid pX01.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP MEDLINE=89172073; PubMed=3148491;
RX Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H.,
RA Schmidt J.J.;
RT "Sequence and analysis of the DNA encoding protective antigen of
RT Bacillus anthracis.";
RL Gene 69:287-300(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=28, 33, BA1024, and BA1035;
RX MEDLINE=99214082; PubMed=10197996;
RA Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;
RT "Genetic diversity in the protective antigen gene of Bacillus
RT anthracis.";
RL J. Bacteriol. 181:2358-2362(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=V770-NP1-R / ATCC 14185;
RX MEDLINE=20359347; PubMed=10899854;
RA Cohen S., Mendelson I., Altobum Z., Kobiler D., Elhanany E., Bino T.,
RA Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,
RA Kronman C., Velan B., Shaferman A.;
RT "Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus
RT anthracis spore vaccines protect against anthrax.";
RL Infect. Immun. 68:4549-4558(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Steirne;
RX MEDLINE=99445483; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Keim P., Koeher T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martinez Y., Ricke D., Svensson R., Jackson P.J.;
RT "Sequence and organization of pX01, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes.";
RL J. Bacteriol. 181:6509-6515(1999).
RN [5]
RP DOMAINS.
RX MEDLINE=91332080; PubMed=1651334;
RA Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;
RT "The carboxyl-terminal end of protective antigen is required for
RT receptor binding and anthrax toxin activity.";
```

Q07833 bacillus su  
P41508 mycoplasma  
P38537 bacillus sp  
P47037 saccharomyc  
Q01550 xenopus lae  
P54697 dictyosteli  
P19321 clostridium  
Q92kw5 helicobacte  
P08964 saccharomyc  
P53968 saccharomyc  
P45949 saccharomyc  
P25389 saccharomyc

34 149 4.1 2334 1 WAPA\_BACSU  
35 148.5 4.1 979 1 P115\_MYCHR  
36 148.5 4.1 1176 1 SLAP\_BACSH  
37 148.5 4.1 1230 1 SMC3\_YEAST  
38 148.5 4.1 1744 1 TANA\_XENLA  
39 148.5 4.1 2245 1 MYSJ\_DICDI  
40 148 4.1 1276 1 BXD\_GLOBO  
41 148 4.1 1288 1 VACA\_HELPJ  
42 148 4.1 1328 1 MYS1\_YEAST  
43 147.5 4.0 678 1 YNC7\_YEAST  
44 147.5 4.0 817 1 YG4A\_YEAST  
45 147 4.0 1037 1 KCC4\_YEAST

RL J. Biol. Chem. 266:15493-15497(1991).  
 RN [6]  
 RP CHARACTERIZATION.  
 RC STRAIN-Sterne;  
 RX MEDLINE=94327640; PubMed=8051159;  
 RA Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;  
 RT "Anthrax protective antigen forms oligomers during intoxication of  
 RT mammalian cells.";  
 RL J. Biol. Chem. 269:20607-20612(1994).  
 RN [7]  
 RP CHARACTERIZATION.  
 RC STRAIN-Sterne;  
 RX MEDLINE=2122592; PubMed=11207581;  
 RA Beauregard K.E., Collier R.J., Swanson J.A.;  
 RT "Proteolytic activation of receptor-bound anthrax protective antigen  
 RT on macrophages promotes its internalization.";  
 RL Cell. Microbiol. 2:251-258(2000).  
 RN [8]  
 RP TOXIN REGULATION.  
 RC STRAIN=Weysbridge;  
 RX MEDLINE=94131936; PubMed=8300513;  
 RA Koehler T.M., Dai Z., Kaufman-Yarbray M.;  
 RT "Regulation of the Bacillus anthracis protective antigen gene: CO2 and  
 RT a trans-acting element activate transcription from one of two  
 RT promoters.";  
 RL J. Bacteriol. 176:586-595(1994).  
 RN [9]  
 RP MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.  
 RC STRAIN-Sterne;  
 RX MEDLINE=95050722; PubMed=7961869;  
 RA Singh Y., Klimpel K.R., Aroa N., Sharma M., Leppla S.H.;  
 RT "The chymotrypsin-sensitive site, PFD315, in anthrax toxin protective  
 RT antigen is required for translocation of lethal factor.";  
 RL J. Biol. Chem. 269:29039-29046(1994).  
 RN [10]  
 RP MUTAGENESIS OF DOMAIN 4 LOOPS.  
 RC STRAIN-Sterne;  
 RX MEDLINE=99185012; PubMed=10085028;  
 RA Varughese M., Teixeira A.V., Liu S., Leppla S.H.;  
 RT "Identification of a receptor-binding region within domain 4 of the  
 RT protective antigen component of anthrax toxin.";  
 RL Infect. Immun. 67:1860-1865(1999).  
 RN [11]  
 RP MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.  
 RC STRAIN-Sterne;  
 RX MEDLINE=21032804; PubMed=11178978;  
 RA Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;  
 RT "Trp 346 and Leu 352 residues in protective antigen are required for  
 RT the expression of anthrax lethal toxin activity.";  
 RL Biochem. Biophys. Res. Commun. 281:186-192(2001).  
 RN [12]  
 RP MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.  
 RC STRAIN-Sterne;  
 RX MEDLINE=21438996; PubMed=11554763;  
 RA Ahuja N., Kumar P., Bhatnagar R.;  
 RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu565, and Ile574 are  
 RT required for oligomerization of anthrax protective antigen.";  
 RL Biochem. Biophys. Res. Commun. 287:542-549(2001).  
 RN [13]  
 RP MUTAGENESIS OF PRO-289.  
 RC STRAIN-Sterne;  
 RX MEDLINE=21255689; PubMed=11356563;  
 RA Khanna H., Chopra A.P., Aroa N., Chaudhry A., Singh Y.;  
 RT "Role of residues constituting the zeta strand of domain II in the  
 RT biological activity of anthrax protective antigen.";  
 RL FEMS Microbiol. Lett. 199:27-31(2001).  
 RN [14]  
 RP MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.  
 EX MEDLINE=21125576; PubMed=11222612;  
 RA Mordridge J., Mourez M., Collier R.J.;  
 RT "Involvement of domain 3 in oligomerization by the protective antigen  
 RT moiety of anthrax toxin.";  
 RL J. Bacteriol. 183:2111-2116(2001).  
 RN [15]  
 RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.  
 RX MEDLINE=21269403; PubMed=11113126;  
 RA Sellman B.R., Nassi S., Collier R.J.;  
 RT "Point mutations in anthrax protective antigen that block  
 RT translocation.";  
 RL J. Biol. Chem. 276:8371-8376(2001).  
 RN [16]  
 RP MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;  
 RX ILE-239; TRP-255 AND PHE-265.  
 RC STRAIN-Sterne;  
 RX MEDLINE=22112896; PubMed=12117959;  
 RA Chauhan V., Bhatnagar R.;  
 RT "Identification of amino acid residues of anthrax protective antigen  
 RT involved in binding with lethal factor.";  
 RL Infect. Immun. 70:4477-4484(2002).  
 RN [17]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 Å STRONGS).  
 RX MEDLINE=97192099; PubMed=9039918;  
 RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;  
 RT "Crystal structure of the anthrax toxin protective antigen.";  
 RL Nature 395:833-838(1997).  
 RN [18]  
 RP REVIEW.  
 RX MEDLINE=21428689; PubMed=11544370;  
 RA Mock M., Fouet A.;  
 RT "Anthrax.";  
 RL Annu. Rev. Microbiol. 55:647-671(2001).  
 CC -1- FUNCTION: One of the three proteins composing the anthrax toxin,  
 CC the agent which infects many mammalian species and that may cause  
 CC death. PA binds to a receptor (ATR) in sensitive eukaryotic  
 CC cells, thereby facilitating the translocation of the enzymatic  
 CC toxin components, edema factor and lethal factor, across the  
 CC target cell membrane. PA associated with LF causes death when  
 CC injected. PA associated with EF produces edema. PA induces  
 CC immunity to infection with anthrax.  
 CC -1- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a  
 CC protective antigen (PA), a lethal factor (LF) and an edema factor  
 CC (EF). None of these is toxic by itself. PA+LF forms the lethal  
 CC toxin (LeTx); PA+EF forms the edema toxin (EdTx). PA-63 forms  
 CC heptamers and this oligomerization is required for LF or EF  
 CC binding. Once activated, at low pH, the heptamer undergoes  
 CC conformational changes and converts from prepore to pore inserted  
 CC in the membrane, forming cation-selective channels.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- DOMAIN: The molecule is folded into four functional domains. Each  
 CC domain is required for a particular step in the toxicity process.  
 CC Domain 1 contains two calcium ions and the proteolytic activation  
 CC site. Cleavage of the PA monomer releases the subdomain 1a, which  
 CC is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is  
 CC part of the remaining 63-kDa fragment (PA63) and contains the  
 CC binding sites for LF and EF. Domain 2 is a beta-barrel core  
 CC containing a large flexible loop that has been implicated in  
 CC membrane insertion and pore formation. There is a chymotrypsin  
 CC cleavage site in this loop that is required for toxicity. Domain 3  
 CC has a hydrophobic patch thought to be a separate domain and shows  
 CC interactions. Domain 4 appears to be a separate domain and shows  
 CC limited contact with the other three domains: it would swing out  
 CC of the way during membrane insertion. It is required for binding  
 CC to the receptor; the small loop is involved in receptor  
 CC recognition.  
 CC -1- PTM: Proteolytic activation by furin or a furin-like protease  
 CC cleaves the protein in two parts, PA-20 and PA-63; the latter is  
 CC the mature protein. The cleavage occurs at the cell surface and  
 CC probably in the serum of infected animals as well; both native and  
 CC cleaved PA are able to bind to the cell receptor. The release of  
 CC PA20 from the remaining receptor-bound PA63 exposes the binding  
 CC site for EF and LF, and promotes oligomerization and  
 CC internalization of the protein.  
 CC -1- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were  
 CC performed that showed that the residues present in the small loop  
 CC of domain 4, and not the ones in the large loop, are involved in  
 CC receptor recognition.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

Query Match 98.8%; Score 3604; DB 1; Length 764;  
Best Local Similarity 96.2%; Pred. No. 3.9e-171;  
Matches 707; Conservative 1; Mismatches 3; Indels 24; Gaps 1;  
QY 1 EVKQENRLNSESSESSQGLLYGFDFLNFQAPMVVTSSTGDLGIPSELENISENQYF 60  
DB 30 EVKQENRLNSESSESSQGLLYGFDFLNFQAPMVVTSSTGDLGIPSELENISENQYF 89  
QY 61 QSAIWSGPIKVKSDPYTFATSADNHVWVDDQEVINKASNSKIRLEKGRLLQIKIY 120  
DB 90 QSAIWSGPIKVKSDPYTFATSADNHVWVDDQEVINKASNSKIRLEKGRLLQIKIY 149  
QY 121 QRENTEKGLDPKLVWTSQNKKEVSSDNILQLPELKOKSSNRKSTSGAPVPRDN 180  
DB 150 QRENTEKGLDPKLVWTSQNKKEVSSDNILQLPELKOKSSNRKSTSGAPVPRDN 209  
QY 181 DGIPDSLEVEGVVDVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEKT 240  
DB 210 DGIPDSLEVEGVVDVKNKRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEKT 269  
QY 241 GRIDKNVSEARHPVAAYPIVHVDWENIILSKNEDQSTQNTDSETRISKNTSTSTHT 300  
DB 270 GRIDKNVSEARHPVAAYPIVHVDWENIILSKNEDQSTQNTDSETRISKNTSTSTHT 329  
QY 301 -----SNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 336  
DB 330 SEVHGNAEVHASFDDIGGSVAGSPSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 389  
QY 337 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQISQILAPNNYPSKNLAPIA 396  
DB 390 NANIRYVNTGTAPIYVLPPTSLVLGKQNTLATIKADENQISQILAPNNYPSKNLAPIA 449  
QY 397 LNAQKXASSTPTIMYNGQFLEKTKQLRLDQVYGNATYFNGVRVDTGNSWSEV 456  
DB 450 LNAQKXASSTPTIMYNGQFLEKTKQLRLDQVYGNATYFNGVRVDTGNSWSEV 509  
QY 457 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTXPDMTLKEALKIAGFNEPKNL 516  
DB 510 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTXPDMTLKEALKIAGFNEPKNL 569  
QY 517 QYQKXDIETFDNFQOQTSQNKQLAELNATYVLDKIKLAKNNILIRDKRFHYDR 576  
DB 570 QYQKXDIETFDNFQOQTSQNKQLAELNATYVLDKIKLAKNNILIRDKRFHYDR 629  
QY 577 NNIAGADESVVKEAREVINSSTEGLLINIDKIRKILSGYVIEDETEGLKEVINDRY 636  
DB 630 NNIAGADESVVKEAREVINSSTEGLLINIDKIRKILSGYVIEDETEGLKEVINDRY 689  
QY 637 DMLNTSSLRODGTFRIDFKKYNKDLPLXISNPKYKVVAVTKNTIINPSENGDTSTNG 696  
DB 690 DMLNTSSLRODGTFRIDFKKYNKDLPLXISNPKYKVVAVTKNTIINPSENGDTSTNG 749  
QY 697 IKKILIFSKKGYEIG 711  
DB 750 IKKILIFSKKGYEIG 764  
RESULT 2  
YPB1\_BACAN  
ID YPB1\_BACAN STANDARD; PRT; 204 AA.  
AC P13422; Q9X377;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein pXOI-111.  
GN pXOI-111.  
OS Bacillus anthracis.  
OC Plasmid pXOI.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.

MEDLINE=89172073; PubMed=3148491;  
Welkos S.L., Lowe J.R., Eden-McCurchan F., Vodkin M., Leppa S.H.,  
Schmidt J.J.;  
RT "Sequence and analysis of the DNA encoding protective antigen of  
Bacillus anthracis.";  
RL Gene 69:287-300(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Storne;  
RX MEDLINE=99445483; PubMed=10515943;  
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,  
RA Kaim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,  
RA Martinez Y., Riecke D., Svensson R., Jackson P.J.;  
RT "Sequence and organization of pXOI, the large Bacillus anthracis  
plasmid harboring the Anthrax toxin genes.";  
RL J. Bacteriol. 181:6509-6515(1999).  
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M22589; AAA22636.1; -;  
DR EMBL; AF065404; AAD32415.1; -;  
DR PIR; G59104; G59104.  
DR PIR; I39933; I39933.  
DR HSP; P3423; IACC.  
KW Hypothetical protein; Plasmid; Transmembrane.  
FT TRANSMEM 162 182 POTENTIAL.  
FT CONFLICT 93 93 T -> A (IN REF. 1).  
FT CONFLICT 184 204 KSCNCILYVEVSQMLNSVFY -> NHVIVYLSM (IN  
FT REF. 1).  
SQ SEQUENCE 204 AA; 23029 MW; E1657B23AB4273FD CRC64;  
Query Match 6.5%; Score 235.5; DB 1; Length 204;  
Best Local Similarity 34.6%; Pred. No. 2.2e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;  
QY 563 MNTILIRKRPYDRNNIAGADESVVKEAREVINSSTEGLLINIDKIRKILSGYVIEI 622  
DB 1 MNTILIRDP-YHYDNGNIGVDDSYLKNAYKQILWSSDGVSLNLDENVQALSQYMLQI 59  
QY 623 EDTE-----GLKEVINDRYDMLNTSSLFQDGTFRIDFKKYNKDLPLXISNPN 669  
DB 60 KKP-SNHLTNSPVTITLAGKDSGVGELYRVLS-----DGTGFLDNKFDENWRSLV-DPG 112  
QY 670 YKYNVAVTKEN-TIINPSENGDTSTNGIKILIFSKKGYEI 710  
DB 113 DDVYVAVTKEDPNAVTRDENGNI-NKLNLTLLVSGKIKEI 153  
RESULT 3  
YM67\_YEAST  
ID YM67\_YEAST STANDARD; PRT; 1658 AA.  
AC Q03661; Q04988;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical 187.1 kDa protein in GUA1-ERG8 intergenic region.  
GN YMR219W OR YMR261.13 OR YMG959.01.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=97313268; PubMed=9169872;  
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,



RA Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S., Jagels K., Lye G., Moulé S., Odell C., Pearson D., Rajandream M.A., Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.",  
RL Nature 387:90-93(1997).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
DR EMBL; Z49809; CA489934.1; -;  
DR EMBL; Z49939; CA90190.1; -;  
DR PIR; S55101; S55101.  
DR Germonline; 142894; -;  
DR SGD; S0004832; ESC1.  
DR GO; GO:0005634; C:nucleus; IDA.  
DR GO; GO:0006348; P:chromatin silencing at telomere; IMP.  
KW Hypothetical protein.  
SQ SEQUENCE 1658 AA; 187137 MW; 3893F968305A757D CRC64;  
  
Query Match 5.2%; Score 188; DB 1; Length 1658;  
Best Local Similarity 19.1%; Pred. No. 0.08;  
Matches 159; Conservative 139; Mismatches 275; Indels 260; Gaps 40;  
  
QY 1 EVKQENLLNESES--SSQGLGYFSDLAFAQPMVTSITGDLSPSELENIPISEN 57  
Db 290 EMELEDDIDVEDAKESQGAEGTEHS-VDFSKYMQPRTDNTKIPVIEKYESDEHKVHQ 348  
QY 58 QYFQSAINS-GPIKVKKSD-----YTFATSAADN---HVTMWVDQEVINKASNSNK 105  
Db 349 RYSEDGAPDFGWNISVDDESEDESEDAESYSANAENVYHNEHLEDDKELIEDIESDS 408  
QY 106 ILEKGRLYQIKIQRENPEKGLDFKL-----YWTDSQNKVEY 145  
Db 409 -----ESQAQESQGSDDFEYKVKNEKSTSEETNTSERSDQGFADYATNKV 459  
QY 146 ISSDNLQLPEL-----KQKSNRKRKSTSGAPTPVRDNDGIPDSLEVE 190  
Db 460 EQQENDEPERKDIIRSLDKNFHNNKSEYSENVLENETDPAIVERNO-IND---VE 515  
QY 191 GTYVDVKNKRTPL-----SPWISNIHE--KKGLTKYKS-----PEKWTASP- 231  
Db 516 GYDVTGKSVESDLHSP--DNLVDLAARAMLQFOQSRNSNCPOKEQVSSYLGHSNGS 573  
QY 232 -----PYSDFEKTGRIDKNVSPFAHPLVAAVPIVHVDMENIILSKNEDQ 277  
Db 574 NLSGRSLDESEIQPLKDF--TGNNNNLKT-D-RGDLSS--VEIEVEKY--SEKK 622  
QY 278 STQNTDSETRTSKNTSTSRHTSNSTVAIDHSLSLAGERTWAETMGINTADTARLN 337  
Db 623 LDGSTKEKELVPLSTDTTINNSLGNDSIYVSLDD-----ADAISENLTD----- 667  
QY 338 AMRVYNTCTAPVNVLPITSLVLGKNTLATIKADENCLS-----QILAP----- 383  
Db 668 --VPLMEKTKTKYEVVISEV-----YSTSYEDNTVAMPQVEYTSFPMNDPNSL 718  
QY 384 NNYFSPKN-----LAPIALNAQKQAS-----STPTIMYNOFLEKTKQL 424  
Db 719 NDDVEKKHDLKSLTALAAPAF--TKDAEFVEAGVTKSLTSTSGHTNIPTSKETKQV 776  
QY 425 RLDTDOVYGNATYVFNGRVVDVGSNWSVLPOIQETIARIIFNGKDLNLVETAAV 484  
Db 777 S-DLDESTENVTFFENGTG-----DENKQSKNFPFGVANSTDKSTEDNTD-----EKYFSAI 827  
QY 485 NPSDFLETTKPDWTLKALKAFGNENPGLQYQCKDITFDFNF--DQQTSON---IKN 540  
Db 828 NVTN-----VTGSSCEDIIETASVVE--NLRYPEKDMNEAEMSSGDCVKQNDGSKT 880

QY 541 QLA-----ELNATNIYTVLDKIKNAKQNTLIRDKRPHYDRNNIAGADESVVKEAHR 593  
Db 881 QISFSTSDPNFOESNDNTEFSSTK-----YKVRNSDLEDESLKELATK 925  
QY 594 -EVIN-----SSTREGLLNIDKDIRKILSGVIVIEITEGKEVI 632  
Db 926 AEVVDKLDDEESEDSEYQDYADPEPQNDGSGNENIVKTKK-----DTLGIVEPE 975  
QY 633 NDRYMLNLSRQDGKTFIDPKYNDKLPYISNPKNYKVNKYAVTKENTIIN 685  
Db 976 NEXVN-----KVHEETLFEANVSSVNVQNKDMHTDVIN 1010  
  
RESULT 4  
MSPL\_PLAFK STANDARD; PRT; 1630 AA.  
ID MSPL\_PLAFK AC P04932;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Merzoite surface protein 1 precursor (Merzoite surface antigens) (PMSEA) (P190).  
GN MSP-1.  
OS Plasmodium falciparum (isolate K1 / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86136024; PubMed=3004972;  
RA Mackay W., Goman M., Bone N., Hyde J.E., Scaife J., Certa U., Stunnenberg H., Bujard H.;  
RT "Polymorphism of the precursor for the major surface antigens of Plasmodium falciparum merzoites: studies at the genetic level."; EMBO J. 4:3823-3829(1985).  
RN [2]  
RP REVISIONS, SEQUENCE FROM N.A.  
RA Pan W., Rolfe R., Bujard H.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (Potential).  
CC -1- PTM: Merzoite surface antigen contain the sequence of 83 kDa, 42 kDa and 19 kDa antigens which are the major surface antigens of merzoites. The maturation take place during schizont.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X03371; CAA27070.1; -;  
DR InterPro; IPR006209; EGF\_like.  
DR Pfam; PF00008; EGF; 1.  
KW Malaria; Merzoite; Polyprotein; Repeat; Signal; Glycoprotein; Transmembrane; GPI-anchor.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.  
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.  
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.  
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).

```
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;

Query Match
Best Local Similarity 4.8%; Score 173.5; DB 1; Length 1630;
Matches 168; Conservative 134; Mismatches 251; Indels 263; Gaps 47;

QY 5 ENRL-LNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENOYFQSA 63
Db 918 ENILSGKNKIYQELIGQKSSSE-NF-----YKILKDSDFTYNE 956
QY 64 IWSGFIKVKSDYEYFATSDAHVMTWDDQEVINKASNS-----NKIRLEKGRLY- 114
Db 957 SFTNFVKSADD-----INSLNDESKRKLEEDINKLKTQLQSFDFLYNKYKLERLFD 1011
QY 115 -----QIKIOYORENPTEKGLDFKLYWTDQNKKEVISSDNLQLPKXKSSNRKKS 168
Db 1012 KKTGVKYNQIKLTLKEQLESKL--NSLNNPKHVL--QNFVSFFNKKKAETATEN 1076
QY 169 TSAGPTVPDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNTHKKGTLTKY---KSSP-- 223
Db 1068 T-----LENTKILLKH-----XGLVKYNGESSPLK 1094
QY 224 ---EKWSTASDPYSDFE---KVTGRIDKNVSEAR-----HPLVAAYPIVHVD 265
Db 1095 TLSEESIQTEDNYASLENFKVLSKLEGLKDNLEKKLSYLSGLHLLIA-----E 1147
QY 266 MENILSKNEDOSTONTSRTISKTSTST-----RHSTNSNSTVAIDHSL 314
Db 1148 LKEVKNK---YTGNSFSENVDNNALESVKFLPREGTVATVVSSEGGDTLEQSQPK 1204
QY 315 SIAGERTWAETGLTADTAPLANIRYNTGTAPIYVNLPTT-----SLVLGRNQTLAT 369
Db 1205 KPASTHVGAE---NTITTSQ-NVDEVDVVIIPFGESEEDYDDLQGVVTEAVTPSV 1260
QY 370 IKADENQLSOLLAPNYPKSNLAPIA---LNAQKDASTTITWYN-----QFLE 419
Db 1261 I---DNILSKI---ENEVEVLKPLAGVYSLKQLENNVMTFNVNKVDILNSRFNKRE 1314
QY 420 KTKQLRLDTPQV-YGNIAI-----YNPENGRVVDVTGNWSEVLPQIQTARIIFN 470
Db 1315 NFKV-LESDDLPIYKDLTSSNVVVDPKYKFNKEKRDKFLSSYVVKDSID----- 1364
QY 471 GKDNLVVERIAVNPSPDL-----ETTKPDM-TLKEALKIAGFNENP-----GNLQY 518
Db 1365 -TDINFA-----NDVLGYKILSEKYSKSDLSISIKKYINDKQGENEKYLPFLNNIET 1414
QY 519 QGKDITE-----FDFNFDQOTSQNIKNQALNATNIYVLDIKLNAXKN- 564
Db 1415 LYKTVNDKIDLPVHLEAKVLNITYEK---SNVEVKIKELN--YKLTQDKLADFKNKN 1469
QY 565 -ILIRDKPHYDRN-----IAGV-ADSEVVKBAHREVINSSTGLLLINIDK--IRKIL 615
Db 1470 FVGIAADLTDYNNHNNLLTKFLSTGMVFENLAKTVLSNLDGNLQ--MLNISQHCQVK-- 1526
QY 616 SGYVIEIDTEGLKEVINDRYDMLNLSISLRDQGTFFIDFKYNDKLPYISNPNKYNVY 675
Db 1527 -----QCQNSGCFPHLDE-----RECKCLNLYKQEGDKC---VENPAPTCH-- 1566
QY 676 AVTKENTINPENG-----DTSNGIKKI 700
Db 1567 -----EN-----NGGCDADAKCTBEDSGSNG-KKI 1590

RESULT 5
MSPL_PLAFW
ID MSPL_PLAFW STANDARD; PRT; 1639 AA.
AC P04933;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Merozoite surface protein 1 precursor (merozoite surface antigens)
DE (PMSEA) (P195).
```

```
GN MSP-1.
OS Plasmodium falciparum (isolate Wellcome).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5848;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86014355; PubMed=2995820;
RA Holger A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,
RA Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,
RA Freeman R.R.;
RT "Primary structure of the precursor to the three major surface
RT antigens of Plasmodium falciparum merozoites."
RL Nature 317:270-273(1985).
RN [2]
RP REVISIONS.
RA Holder A.A.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor
CC (potential).
CC -2- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42
CC kDa and 19 kDa antigens which are the major surface antigens of
CC merozoites. The maturation take place during schizont.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; X02919; CAA36676.1; -.
DR PIR; A24594; A24594.
DR PIR; S05603; S05603.
DR PDB; 1CEB; 28-MAY-99.
DR InterPro; IPR006209; EGF_like.
DR Pfam; PF00008; EGF; 1.
KW Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 1639 MEROZOITE SURFACE PROTEIN 1.
FT CARBOHYD 116 116 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 768 768 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 783 783 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 844 844 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 920 920 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;

Query Match
Best Local Similarity 4.8%; Score 173.5; DB 1; Length 1639;
Matches 168; Conservative 134; Mismatches 251; Indels 263; Gaps 47;

QY 5 ENRL-LNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENOYFQSA 63
Db 927 ENILSGKNKIYQELIGQKSSSE-NF-----YKILKDSDFTYNE 965
QY 64 IWSGFIKVKSDYEYFATSDAHVMTWDDQEVINKASNS-----NKIRLEKGRLY- 114
Db 966 SFTNFVKSADD-----INSLNDESKRKLEEDINKLKTQLQSFDFLYNKYKLERLFD 1020
QY 115 -----QIKIOYORENPTEKGLDFKLYWTDQNKKEVISSDNLQLPKXKSSNRKKS 168
Db 1021 KKTGVKYNQIKLTLKEQLESKL--NSLNNPKHVL--QNFVSFFNKKKAETATEN 1076
QY 169 TSAGPTVPDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNTHKKGTLTKY---KSSP-- 223
```

Db 1077 T-----LNTKLLKHV-----KGLVYKNGESSPLK 1103  
Qy 224 ---EKWSTASDPYSDFE-----KVTGRIDKXVSPEAR-----HPLVAAPVHVVD 265  
Db 1104 TLSESTQIOTEDYASLENPKVLSKLEGLKONLNLEKKKLSYLSGLHLLIA-----E 1156  
Qy 266 MENIILSKNEDQSTQNTSETTISKNTSTS-----RHTSNSTSTVAIDHSL 314  
Db 1157 LKAEVIKKNK---YTGNSPENNNDVYNALESYKXKLPGSTVATVVSSESGDTLEQSQPK 1213  
Qy 315 SLAGERTWAETMGLNTADTARLANAIRVYNTGTAPIYVLPFT-----SLVLGKNQTLAT 369  
Db 1214 KPAETHVGAES---NTITTSQ-NVDEVDVDDVLIIVFGESEEDYDQGVVGTGEATPSV 1269  
Qy 370 IKADENQLSQILAPNYPSPKSLAPIA---LNAQKQDASSTPTMYN-----OFLELE 419  
Db 1270 I---DNILSKI---ENEYEVLYKLPGVYRSLSKQLENNVMTFVNVKDIILNSFENKRE 1323  
Qy 420 KYQQLRLDQV-YGNIAT-----YFENGVRVDTGNSWSEVLPOIQTETARIIFN 470  
Db 1324 NFKNV-LESDLIPYKDLTSSVYKDPYKFLNKEKDKFLSSYNIKDSID----- 1373  
Qy 471 GQDLNLVERRIAAVNPSPDL-----ETTKPDM-TLKEALKIARFGENEPN-----GNLQY 518  
Db 1374 -TDINFA-----NDVLGYKILSEYKSDLSIKKYINDKQGENEKYLPFLNNIET 1423  
Qy 519 QOKQITE-----PDFNEQOTSQNIKNQLAELNATNTYTVLDKIKLNAGVN- 564  
Db 1424 LYKTVNDKIDLFVHLEAKVLNVTYVEK---SNVEVKIKELN---YLKTIQDKLADPKKNN 1478  
Qy 565 -LLIRKPRHYDRNN---IAVG-ADESQVKEAREVINSSTEGILLNIDKO---IRKIL 615  
Db 1479 FVGIADLSTDYHNHLLTFLSTGVFENLAKTV-SNLDGMLQG-MNISQHQCVK- 1535  
Qy 616 SGYVIEBTEGLKEVINDRYDMLNISSRQDKTFIDPKYNDKPLIYNPNYKNVY 675  
Db 1536 ---QCPQNSQCFRLDE-----RECKCLLNKYQEGDKC---VENPNPTCN-- 1575  
Qy 676 AVTKENTINPSENG-----DTSNGIKKI 700  
Db 1576 ---EN-----NGCDDAKCTEEDSGSNG-KKI 1599

RESULT 6

PCP1\_SCHPO STANDARD; PRT; 1208 AA.  
AC Q235I; 1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Spindle pole body protein pcpl.  
GN PCP1 OR SPAC6G9.06C.  
CS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.  
RX MEDLINE=21852775; PubMed=11864908;  
RA Flory M.R., Morpew M., Joseph J.D., Means A.R., Davis T.N.;  
RT "pcpl, a Spcl10p-related calmodulin target at the centrosome of the  
RL fission yeast Schizosaccharomyces pombe";  
RN Cell Growth Differ. 13:47-58 (2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Scouras J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
RA Borzom K., Langer I., Beck A., Lehmach R., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armetrong J., Forburg S.L.,  
RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Sipakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe";  
RL Nature 415:871-880 (2002).  
CC -I- FUNCTION: Spindle pole body component that binds calmodulin.  
CC Overexpression of pcpl causes the formation of supernumerary SPB-  
CC like structures and disrupts both mitotic spindle assembly and  
CC chromosome segregation.  
CC -I- SUBCELLULAR LOCATION: Spindle pole body.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL; Z81317; CAB03608.1; -;  
DR EMBL; AF348506; AAK31344.1; -;  
DR PIR; T39068; T39068.  
DR GeneDB SPombe; SPAC6G9.06C; -;  
KW Calmodulin-binding; Coiled coil.  
FT DOMAIN 151 375 COILED COIL (POTENTIAL).  
FT DOMAIN 387 803 COILED COIL (POTENTIAL).  
FT DOMAIN 874 1091 COILED COIL (POTENTIAL).  
FT DOMAIN 1177 1204 COILED COIL (POTENTIAL).  
SQ SEQUENCE 1208 AA; 140763 MW; 70264159ADD42424 CRC64;  
Query Match 4.7%; Score 171.5; DB 1; Length 1208;  
Best Local Similarity 20.0%; Pred. No. 0.34;  
Matches 151; Conservative 132; Mismatches 328; Indels 145; Gaps 31;  
Qy 10 NESSESSQGLLG-----YVFSLNFOAPMVVTSSTTGLSPSSSELENIPSPNQYF 60  
Db 78 NKNSDKYNGSLGDKGSPDPNSYGLSALKQATQCEALSISQGNDSYDVSKLTLSKNSKI- 136  
Qy 61 QSAIWSGFIVKKSDEYTFATSNADNHVTVWVDQEVINKASNSN-----KI-----RLEK 110  
Db 137 -----DHTDGLPANAALTIREQEVLEKVSRENFGRLIKIVCLEKRLS 181  
Qy 111 GRLVQIK-----IQYQRENPTKGLDFKLTWYDTSQNKKEVISSDNLQPLKQKSSN 162  
Db 182 MAEQIKAEVKNVDELHAERAN-----LQIQKRTES-----LLQKSED 220  
Qy 163 SRKRSTSAQFTVDRNDGIDPSLEVEGYTVDKVKRTFLSPWISNIHKKGLTKYKSS 222  
Db 221 KNFKLEKV--DYLKSKVND-VEQSNVQVFTERIR-----FLENALKVQREKDSLSTEME 273  
Qy 223 PEKWTASDPYSDFEYKVTGRIDKNVSPFAHPLVAAPYIVHVDMENTILSKNEDQSTQNT 282  
Db 274 EDKSNKEVDYEYTRQQLNRLDE-LSEE-----LDVAQDLTEKDETA--- 316  
Qy 283 DSETRTISKNTSTSRHTSNSTNSSTVAI--DHSLSLAGERTWAETMGLNTADTARLAN- 339

Db 317 -TLXQRIEKENSSAFENBESSVYHLQEDYAILQAKCDEFADRIQVLTADLEKENQ 375  
QY 340 IRYNTGTAPIYVLPPTSLVGNKQTLATIKADENQSLQILAPNNYPSKNLAPIALNA 399  
Db 376 IMHSEASIGTDSMOVHTLOEQHKAHEEFLDQISRMNEEGKNFEDIMLQFSLSE 435  
QY 400 QKDASSPTITWYNOFLEKTKRLDQDOVGNATYFNGRVRVDTGNSWSEVLPQ 459  
Db 436 ERDVLESKL-----QTLE-DNNSRLMTSSLGNQIESLQNR--EIDEEKHLRLAS 487  
QY 460 IOETTARIIFNGKOLNVERIAAVNSDPLETTKPDWTLKBAKIAFGNENPNQYQ 519  
Db 488 -----KNSDKALAEATNRLQEVTKLETLR--MKNNDLNEIHDRENEGLTK 535  
QY 520 GKDITEFDNFQDQTSNKK--QLAELNATNITVTLDKIK-----LNAMNIL-IRD- 569  
Db 536 IDSITKEKORLINELEQRIKSYEVNSELNGT-IDEYRNKLKXKEEYNEWMNAFOYKN 594  
QY 570 --KRFPHYDRNNAVGADESVVKEARHEVINSSTGLLLNIDKO---IRKILSGYVIEID 624  
Db 595 DLRRFHESINKLQREKELTSLNKKNLVLSLRETVMLEKEREISKYLSGNNAKDLN 654  
QY 625 TEGLEKVINDRY-----DMLNITSLRQDGKTFIDPKYNDKLPVINSNPYKN 673  
Db 655 T-NLMEILNDKISVLQRLTDVKDELQVSEERE-EAIVAGQKUSASFEL-MSNEKQALE 711  
QY 674 VYAVTKENTIIIPNSGDTSTNGIKKILIFSKGYE 709  
Db 712 LKYSKLNELINAQNLDRREELSEL---SKLFE 744

RESULT 7  
YJL3 YEAST  
ID YJL3 YEAST STANDARD; PRT; 1803 AA.  
AC P47024; P87192;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Transposon Ty4 207.7 kDa hypothetical protein.  
GN Ty4 OR YJL113W OR J0780.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / FY1679;  
RX MEDLINE=97103775; PubMed=8948101;  
RA Cziepluch C., Kordes E., Pujol A., Jauniaux J.-C.;  
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X  
RT reveals 19 open reading frames including URA2 [5' end], TRK1, PBS2,  
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CTF7, GZF3, two tRNA genes,  
RT three remnant delta elements and a Ty4 transposon.";  
RL Yeast 12:1471-1474(1996).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z49389; CAA89409.1; --  
DR SGD; S0003649; YJL113W.  
DR InterPro; IPR001584; Rve.  
DR InterPro; IPR001878; Znf\_CCHC.  
DR Pfam; PF00665; rve; 1.  
DR SMART; SM00343; Znf\_C2HC; 1.  
KW Transposable element; Hypothetical protein.  
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;

Query Match 4.5%; Score 163; DB 1; Length 1803;

Best Local Similarity 19.6%; Pred. No. 1.5;  
Matches 161; Conservative 118; Mismatches 262; Indels 282; Gaps 41;  
QY 1 EVKQENRLINSESSSQGLLYFSDLNFPQPMVVTSTTGDLSIPSELENISENOCYF 60  
Db 1078 QLKTN--HETSPFKESIG--TNVFRNTNNEISLKTGTSPFKTLESINHH-- 1128  
QY 61 QSAIWSGFIVKVKSDYTFATSNADNHVMTWDDQEVINKASNKIRLEKGRLYQIKY 120  
Db 1129 -----SDNY-----STNKVE-----KF 1140  
QY 121 QREN-----PTEKGLDFKLYWTDSONKKEVISDNQLPDLKQKSNRSKSTSGPTVP 176  
Db 1141 EXENHHPPIEDIVM-----SDQTDMSNCQDGNLKLKVTDKNVPDTGNTNSPR-- 1193  
QY 177 DRDNDGIDPSLVEGVYDVVKNRFTLSPWISNIHEKK-----GLTKY----- 219  
Db 1194 -----LEONIASPSPQTVNKSFLNKEFSSLNKKRKHEDKNSITSYELERDKR 1247  
QY 220 -----KSPKSWSTASDPYSDFEKTGRIDKNQSPPEARHPLVAAYPIVHVDMENILSK 273  
Db 1248 SKNRVKLIPDNMTVSAPKIRAIYNEAISKNPDLKEKHEYKOAY--HKELQNLKDKM 1304  
QY 274 NEDQSTQNTDSE-----TRT-----ISKNTSTSRHTSNSNS---STVAIDHS 313  
Db 1305 VFDVDVKYSRSRSEIPDNLIVPTNTITFKKNGIYKARIVCRGDTQSPDTYSVITTESLHN 1364  
QY 314 -----LSLAGERT-WAETMGLNTADT-ARLNANIRIVNTGTAPIYVNLVLP-TTSLVLGNQX 365  
Db 1365 HKIFLMIANNRNMFMKTLIDINHAFLYAKLEBE-----IYIPHPRRCVVKLNK 1414  
QY 366 TLATIKADENQ-----LSOI-LAPNNYB-----SKNLAPIALNAQKSDSTPTM 410  
Db 1415 ALYGLKQSPKENDHLRQVNLGIGLKDNSYTPGLVQTEDKNLM-IATVDDCVIAASNEQ 1473  
QY 411 NYNQFL-ELEKTKQLR-----LDTQVYGNATYNFENGRV-----RVDTGS 451  
Db 1474 RLDEFINKLSNFELKITGTLDDVLDTD-ILGMDLVNKRGLGTLDTLTKSPINRMD--K 1530  
QY 452 NMSEVLPOQTETARIIFNGKOLNVERIAAVNP-SDPLETTKPD-----MTLKEALKI 505  
Db 1531 KYNEELKIRKSIPIHSMSTYK-----IDPKQDVLMSEBEFFRGVLKQQLL-- 1577  
QY 506 AFGNEPNGNLQVQGDITEFDNFQDQTSQNIKNQLAELNATNITVTLDKLNAKMMI 565  
Db 1578 -----GELNVV-RHKCRYDIEFAVKVAVLVNYPHERVFMVYKIIQYL-----V 1621  
QY 566 LIRDKRFHYDRN-----NTAVGADESVVKEAHE-----VI-----NSSTEGLL 604  
Db 1622 RYKDIGIHYDRDCNKDKKVIATDASVGSSEYDAQSRIGVILWYGMNIFNVYGNKSTNRCV 1681  
QY 605 LNIKDIDIRILSGY-----IVEIEDT-----EGLK----- 629  
Db 1682 SSTEAEHALYEGYADSEFLKVLKELGSDNNDIVMTDSKPAIQGLNRSYQQPKKFT 1741  
QY 630 -----EVINDRYDMLNTSSLRQDGKTFI-----DFKKY 657  
Db 1742 WIKTEIIEKIKEKSIKLLKITGKNIAIDLTKKPVASDFKRF 1784

RESULT 8  
RIP3 MOUSE  
ID RIP3 MOUSE STANDARD; PRT; 1024 AA.  
AC P57434;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Rho-interacting protein 3 (p116RIP) (RIP3).  
GN RHOIP3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;

[1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=97344280; PubMed=9199174;  
RA Gebbink M.F.B.G., Kranenburg O., Poland M., van Horck F.P.G.,  
RA Housa B., Moolenaar W.H.;  
RT Identification of a novel, putative Rho-specific GDP/GTP exchange  
RT factor and a RhoA-binding protein: control of neuronal morphology.;  
RL J. Cell Biol. 137:1603-1613(1997).  
CC FUNCTION: RhoA-binding protein involved in control of the actin  
CC cytoskeleton. Overexpression promotes neuronal cell flattening and  
CC neurite outgrowth, probably by counteracting rhoA-mediated  
CC signaling.  
CC TISSUE SPECIFICITY: Highly enriched in the brain.  
CC SIMILARITY: Contains 2 PH domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U73200; AAB18198.1; -  
CC FIR; T30869; T30868.  
CC MGD; MGI:1349438; Rhoip3.  
CC InterPro; IPR001849; PH.  
CC Pfam; PF00169; PH; 2.  
CC SMART; SM00233; PH; 2.  
CC PROSITE; PS50003; PH DOMAIN; 2.  
KW Guanine-nucleotide releasing factor; Repeat; Coiled coil.  
FT DOMAIN 43 150  
FT DOMAIN 386 482  
FT DOMAIN 164 172 PRO-RICH.  
FT DOMAIN 185 190 POLY-SER.  
FT DOMAIN 285 296 PRO-RICH.  
FT DOMAIN 671 976 COILED COIL (POTENTIAL).  
SQ SEQUENCE 1024 AA; 116363 MW; A918DBB955245E02 CRC64;  
  
Query Match 4.5%; Score 162.5; DB 1; Length 1024;  
Best Local Similarity 19.2%; Pred. No. 0.77;  
Matches 157; Conservative 131; Mismatches 304; Indels 225; Gaps 39;  
  
QY 31 APWVTSSTG---DLSPSSSELENIPSENYFOSAIWSGFIKVKSDVTFATSDNH 86  
DB 173 AKWAVTSSGGTSGSSSSIPSA-EKVPTT---KSTLWQEMRAKDPDGTSLSPAQSP 226  
QY 87 VTWW-----VDDQEV---NKASNSKIRLEK---RLYQIKIYORENPTEK 128  
DB 227 SQSQPPAACTPREPGLESKEDESTISGDRVDGGRKVRVESGYFLEKAKQLRAEQLPP 286  
QY 129 GLDFKLYWTDSONKQEVSS-DNLQLPE-----LKQKSSNSRKRS----- 168  
DB 287 LLSPPSPSTPHRRSQVIEKFEALDIEKAHMETNMLILTPSSDTQGRERRAIPKR 346  
QY 169 --TSAGTVPDRNDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKG-LTK-YKSSPE 224  
DB 347 DPASEAPTAP--LSDACPLSPHRRAKSLDRRTSSMTPLDINF--KKGWLTQYEDGQW 402  
QY 225 K--WSTASDP---YSD-FEKTGRIDKNVSPEARHPLVAAPYIVHVDMENILSKNEDQ 277  
DB 403 KKHFWFLADQSRUYRDSVAEAAADLGEINLSTCYD-VTEYPV-----QRNYGF 451  
QY 278 STONTSETRTISKNTSTSRTHTSNSSTVAIDHSLSLAGERTWAETMGLNTADTARLN 337  
DB 452 QIHTKEGEF-TLSAMTSGIR-----RNWQIIT----- 477  
QY 338 ANIRYVNTGTAPYVLPITTSVLGNQTLATIKADENQSLAPNPNYPSKNLAPAL 397  
DB 478 --MKHYLPASAPDV---TSSLPEGNKGTSPETCSRSTEQEABPGEPDPEQKKSARE 531  
QY 398 NAQKDSAST-----PITWNYNQ-----FLELEKTKQ----- 423

DB 532 RRREGSKTFDWAEPFPIQALQASAVGSSDSGDCGCLAEAPGELERARRRREPR 591  
QY 424 -----LRLDTDQVYNIATYFENGVRVDTGSNWSEVLPIQIETT- 464  
DB 592 KRFGMLDTIDGFMEDTALRMDIDRSPGLGTPDPUKQNVHVEIQRWQV-----ETTP 646  
QY 465 -----ARIIFNGKD-----INLVERRI--AAVNPSPLETTTPDMLTKAL 503  
DB 647 LREKQVPIAPLHLSLEDRSERLSLTHLSLEKEQSQKEASDLLBNR---LLQDQL 703  
QY 504 KIAFGENENGMLQGGKIDITFDNFQOQTSQNIKNQLAELNATNITVLDKIKLNAM 563  
DB 704 RVALG-REQSAREGYVQATCGRFAAMEETHQKIEDLQROHQRELE-----KREK 756  
QY 564 NILIRDKRPHYDRNNIAGADESVVKEAHE-----VINSSTEGLLNIDKDIRKILSGYI 619  
DB 757 DLLAAEE-----TAATISAIEA-MKNAREHREMERLEKSQRSQISSINDIEALRRQYL 809  
QY 620 VEIETDEGUKVINRYDMLNT-----SSLQDQKTFIDFKYNDKPLIYISNPNKYN 673  
DB 810 EELQSVQRELEVLSEYQSKLENLAHALAQAERQALRQCORENQELNAHQELNNRL- 868  
QY 674 VVAVTKENTIIINPSENGDTSTNGIKKILIFSCKGYEI 710  
DB 869 AABITRLTLTGCGGEGST--GLP--LTQGKDAYEL 901  
  
RESULT 9  
ID YF16 YEAST STANDARD; PRT; 1233 AA.  
AC P43537;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Hypothetical 137.7 kDa protein in UGSI-FAB1 intergenic region.  
GN YF016C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=95400292; PubMed=7670463;  
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
RA Yamazaki M., Tashiro H., Eki T.;  
RA "Analysis of the nucleotide sequence of chromosome VI from  
RT Saccharomyces cerevisiae.";  
RL Nat. Genet. 10:261-268(1995).  
CC SIMILARITY: SOME, TO MAMMALIAN NEUROFILAMENT TRIPLET M PROTEIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D50617; BAA09255.1; -  
CC FIR; S56271; S56271.  
CC GerMOnline; 140170; -  
CC SGD; S0001912; YF016C.  
CC Hypothetical protein.  
KW SEQUENCE 1233 AA; 137697 MW; C8A7CD2C6F0892F6 CRC64;  
  
Query Match 4.5%; Score 162.5; DB 1; Length 1233;  
Best Local Similarity 18.9%; Pred. No. 0.98;  
Matches 151; Conservative 111; Mismatches 253; Indels 285; Gaps 34;  
  
QY 40 TGDLSI---PSSELENIPSENYFQSAIWSGFIKVKSDVTFATSDNHTVWDDQEV 96

```
Db 458 TNDISVEKEEEEEEEEN-----STFSVKK-----ENVT---GEQEA 495
Qy 97 I--NKASNSKIRLEKGRLYQIKIQORENTEKGLDFKLYWTDSONKKEVIGSDNLQLP 154
Db 456 VRNEVSTCEESTSKGE---EIMGDEKSEAGEKSSIIEISGSANSKISKNLVLE 551
Qy 155 ELKQKSNRKRSTSAFTPPDRNDG-----IPDSLEV----- 189
Db 552 DEAEPTQENKFTVEVIGSIDPAPRDDVDVEIVEAVERKNIIPEDELEVAKEQGEQVKLDE 611
Qy 190 -----EGYTVVQNKRTFLSPWISNIHEKGL-----TKY 219
Db 612 PVKAMKDKIAMRGAESISEDKKQEGTAB-LSNEKAKKEVDETABESAGVEVEKST 670
Qy 220 KSSPE--KWTASDPYSFEKVTGRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQ 277
Db 671 PESPRVKRCTSGRP-EDLQ-INERDPEVLKEDVRVPDEDVKPIATTIEN---SEBEDP 725
Qy 278 STQNTDSTRITSKNTSRTHTSNSSTVAIDH-----SLSLAGER-TWASTMGINT 330
Db 726 KSQROIQT---EQAEITQDMGVGTTTFKBEKPRFEITQEGDKITGKDTNHEHG 781
Qy 331 ADTARLANIRYNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNYYPSK 390
Db 782 EATEAASENSKASDVGTAEKY-IEPSSSV--KKOTEDEAVENSEKTEFI-----K 830
Qy 391 NLAPIA-LNAOKDASSTPTWYNQFLEL-----EKTKQLRLDTPQVYGNI--- 435
Db 831 VKAELENLDAPEAEVTAELNKENEDVEVDTEDEAVENSEKTEFIKVKAB--LGNLDAP 888
Qy 436 -----ATYFNGVRVVDTSNWSGVLPQIQETTARIIFNG-----K 472
Db 889 KEAEVTAELNKENEDVEVAATSK-EDIETKSEPAETPIEDTCTAEVSKKDAEAVTKE 947
Qy 473 DLNLVERRTAA-----VNPSDPL----- 490
Db 948 DENMENSKEAEALQVGTQDEIDDDNISDEFQRTVELPELEKQDKONKGEKLEVEET 1007
Qy 491 --ETTKPMDLKEALKIAPGFNEPNQLQYQKGDIETFDNF-----FD 531
Db 1008 EKETSPLDLVVEE-----NITEKNEIKQEEBVEYSQLDNFNETESISKEAPNNDENGPEF 1061
Qy 532 QOTSQNIKNQLAB-----LNATNITYVLDKIKLNAKNILIRDKRHYDRNNIATGADE 585
Db 1062 QSTRENPKASADDIFKQILDDETNEFLQKLVDDSELNALLQ-----SLDAK 1110
Qy 586 SVVKEAAREVINSSTGLLLNTDKDIRKILSGVIVEIEDTEGLKEVINDRYDMLNLSLR 645
Db 1111 STQ-----TTEQSKNDKP----- 1126
Qy 646 QDKTFFDKYNDKLPYI 665
Db 1127 QDVITTSIRKLNKEPVYI 1146

RESULT 10
RBP2_PLAVB STANDARD; PRT; 2867 AA.
AC Q00739; Q9NZM3;
DT 01-APR-1993 (Rel. 25, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Reticulocyte binding protein 2 precursor (PVRBP-2).
GN RBP-2 OR RBP2.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A., AND REVISIONS TO 2438-2439.
RX MEDLINE=20239192; PubMed=10838229;
RA Galinski M.R.; Xu M.; Barnwell J.W.;
RT "Plasmodium vivax reticulocyte binding protein-2 (PVRBP-2) shares
```

```
RT structural features with PVRBP-1 and the Plasmodium yoelii 235 kDa
RL thophry protein family.";
RN Mol. Biochem. Parasitol. 108:257-262(2000).
RP [2]
RX SEQUENCE OF 1189-2439 FROM N.A.
RA MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R.; Medina C.C.; Ingravallo P.; Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
RL merozoites.";
RL Cell 69:1213-1226(1992).
CC -!- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
CC human reticulocyte cells.
CC -!- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF184623; AAF7525.1; -.
DR HSP; P03069; IGCW.
KW Malaria; Receptor; Signal; Transmembrane; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 2867 RETICULOCYTE BINDING PROTEIN 2.
FT DOMAIN 22 2805 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2805 2826 POTENTIAL.
FT DOMAIN 2827 2867 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 44 133 ASN-RICH.
FT DOMAIN 560 758 LYS-RICH.
FT DOMAIN 1112 1285 LYS-RICH.
FT DOMAIN 2758 2785 7 X 4 AA TANDEM REPEATS OF H-D-D-T.
FT REPEAT 2758 2761 1.
FT REPEAT 2762 2765 2.
FT REPEAT 2766 2769 3.
FT REPEAT 2770 2773 4.
FT REPEAT 2774 2777 5.
FT REPEAT 2778 2781 6.
FT REPEAT 2782 2785 7.
SQ SEQUENCE 2867 AA; 331433 MW; 6E7D8CA71AFBFD3 CRC64;
```

```
Query Match 4.4%; Score 161; DB 1; Length 2867;
Best Local Similarity 18.4%; Pred. No. 3.6;
Matches 162; Conservative 147; Mismatches 302; Indels 268; Gaps 42;

Qy 9 LNESESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENYQFQSAIWS-- 66
Db 927 LNELESG-----FKSLN-----GSASTTNKQENIRKNTDTIKSLNPAKN 966

Qy 67 -----GFTK-----VKSDVETPATSDNHVTMWVDDQF-----VINKASNSK 105
Db 967 SSSSKLSLENIINKKADLIKLDQHT--QEIEKHT--FIENEEMSPLSVIKKEKNRVE 1022

Qy 106 IRLEKGRLYQIK-----IOYQR-----ENPTEKGLDFKLYWTDG----- 139
Db 1023 SDMSSELIKQNTKINALEYVYNSKDRFNGDDDETNLLELDDFKQCOQAQOIKKLTN 1082

Qy 140 -----ONKKEVISSDN-----LQPELKQKSSNSRKRSTSAFTPPDRD 179
Db 1083 YNVLDNGINVIIEKEHEKVIILSENHITKDKKINEKIQNVNSLNMETKGLGLKINED 1142

Qy 180 NDGIPDLSL---EVEGY-----TVDVKNKRTFLSPWISNTHKEKGLTKYSSSPKWS 227
Db 1143 IKNSRDTTIKSKIQEFKKVQTIFGSDIVANKK-----IDAKKHVDYNDKDFDEKVK 1196

Qy 228 TAS--DPYSDPEKVTGRI-----DKNVSPEARHPLVAAYPIVHVD----- 265
Db 1197 DTSPEKKKGIKAYEKGKNTLKEKMDDEKNIKEVEBAQI-QYKRIFFIDHVNLMND 1255

Qy 266 -----MENILISKNE-DQSTQNTDSETRTISKNTSTRTSTNSNST---VAIDHS 313
```

```

Db 1256 EVEKSKIVMEKIELYKKEIDEIKQKTE-----YKQDTSNFYTYEQVNSATQSKAKIEQF 1311
Qy 314 LSLAGERTWAETMGLTADTALNANIRVNTGTAPIYVNLPTTSLVGLKQKQTLATIKAD 373
Db 1312 INIA-----TTKGTSTSQ-----DINELESIKEEVHKH--LQLVQOE 1348
Qy 374 ENQL-----SQILAPNNYPSKLNAPIALNAQKADASTPTMTNNOFLELTKQLRLDTD 429
Db 1349 SNSMEEMRKQILSMKDLLLNSEETIA---KEISNN--TQNALGFRENKTLNK--TD 1400
Qy 430 QYVGNATYFNGRV---RVTGNSWSEVLQIQ--ETTAIIIFNGKDLNLVERRIAIV 484
Db 1401 ELLQRYAAM--IBEAKAHKNIDIALEDQAIDTEVSKIEQINREIMNKXD--EIKSYLSEI 1457
Qy 485 NPSDPLETTKPMWTKEALKIAF-----GFNEPNGNLOVQGGKDIPEPDFN 529
Db 1458 KEYKDKCTTEISNKGKDKIEFLEKFKENESNKNVINEINIRNSOYLKDIB-- 1515
Qy 530 FPOQTSQNTQKQALNATNIVTVLDKIKLNKKNILIRDRKFRHYDRNIAVGADESVMK 589
Db 1516 -DAEQKASTKVELFKHETTSINIFKESI---LGVETKSQK-----KINKAEDIMK 1563
Qy 590 --EAREVINSTEGLLNIDDKIRKILSGYIVIEIDTEGLKEVIND----- 634
Db 1564 EIERHNSIEQTQVKGFEQENLANK-----LNEPHYDNDAEDELNDKSTNAKVLITN 1614
Qy 635 ----RYDMLNISLRQDG-KTFIDFKYNDKULPLYISN-----PNYKVNYYA 676
Db 1615 LESVKNLSEITNIKQGGKIKYKAKDIMOKIKATSENTAEKTLBKVDQDSNYVYLNQ 1674
Qy 677 VTKENTIIIPSENG---DTSNGIKKILIFSKGVEIG 711
Db 1675 ITERNLIIVTEKNRLNGIDSTTINIEGALKESKGNVEIG 1713

```

## RESULT 11

```

MDS3 YEAST
ID MDS3 YEAST STANDARD; PRT; 1487 AA.
AC P53094;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE MDS3 protein (MCK1 dosage suppressor 3).
GN MDS3 OR YGL197W OR G1307.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_taxid=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97197971; PubMed=9046087;
RA Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P.,
RA Bruschi C.V.,
RT "Sequencing of a 40.5 kb fragment located on the left arm of
RT chromosome VII from Saccharomyces cerevisiae.";
RL Yeast 13:55-64(1997).
RN [2]
CHARACTERIZATION.
RP Li M.B., Neigeborn L.;
RA Unpublished observations (XX-1997).
RL
CC -!- FUNCTION: NOT KNOWN; NEGATIVE REGULATOR OF EARLY MEIOTIC GENE
CC EXPRESSION.
CC -!- SIMILARITY: CONTAINS 3 KELCH REPEATS.
CC -!- SIMILARITY: TO YEAST YER132C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC
DR EMBL; X91837; CAA62947.1; -
DR EMBL; Z72719; CAA96909.1; -
DR PIR; S62048; S62048.
DR GenOnline; 141245; -
DR TRANSFAC; T03484; -
DR SGD; S0003165; MDS3.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0030435; P:sporulation; IGI.
DR InterPro; IPR006652; Kelch_rep.
DR Pfam; PF01344; Kelch; 2.
KW Kelch repeat; Repeat; Meiosis.
FT REPEAT 171 226
FT REPEAT 234 287
FT REPEAT 371 419
FT REPEAT 434 484
SQ SEQUENCE 1487 AA; 167073 MW; 768AEFBAB796E447 CRC64;

Query Match 4.4%; Score 160.5; DB 1; Length 1487;
Best Local Similarity 19.3%; Pred. No. 1.6;
Matches 171; Conservative 137; Mismatches 271; Indels 305; Gaps 47;

Qy 2 VKOENLLNSESSESSQGLLYGFSDLNFOAPMVVTSS-----TTGDLISIPSSSELENIPS 55
Db 610 VRGEGE-----SSGSLDNVF-EKNF--PIPARTSVSEAQNTQOVANAKAENTPS 660
Qy 56 ENQYFOSALMSGFIKVKSGDEYTFATSDNHYTMV-----DDQEVIN--KASNSKI 106
Db 661 -----TSDEPSSSSDLYSTPHYORNDEBDDPVPSPKPVSKNSI 703
Qy 107 RLEKGLYQ--IKIQORENPTEKGLDFKLYWTD-----ONKKEVSSDNLQPELKQKSS 161
Db 704 -----YPIRKTSSSTSSNGMIFRVPFKEKAATVTEALLESLSLOELSRRS 756
Qy 162 N-----SRKRSTAGPTVPRDNDGIP-----DSLEVEGYTVDVKNKT 201
Db 757 SLMSIPSGELLRSSISEAEHQRASHPLTSSPLFEDSGTFCGKQLQLOQHT--IQPHN 814
Qy 202 FLSP-----WISNIHEKGLTKYKSPKKNSTASPYDFEKVGTGIDKNVSP 249
Db 815 HLSFRFRSRSARSISYVSSSSDRRG-----NSISRSST-SDSFGT-PVLGLVNLVPLPP 867
Qy 250 EARHPLVAAPYIVHVDMENIILSKNDQ---STQNTDSETRITISKNTSRTSRHTNSN-- 304
Db 868 QREP-----NEPPPCPAMSGSNTR--RSNLTLDYVHNSKASFP 906
Qy 305 -----SSTVAIDHSLAGERTWAETMGLTADTALNANIRVNTG--TAPIY 351
Db 907 SRRSSRHIGRRSTPETENAFS--ATPRA-----SLDQMLGKSLKEGSTSQYTQPRM 957
Qy 352 NVLPTTSLVLGKQTLATIKADEN-----QLSQILAPNNY- 387
Db 958 NSFPKA-----NETIQTSSNNESRQSVTSNTDPSLSQSNFALELEPLITFSLYM 1011
Qy 388 --PSKN-----LAPIALNAQKADASTPTMTNNOFLEL-----EKT 421
Db 1012 PWTSTVTRAPAEFFYTGQVNSKWLAPVALDLVAKIYEIPLLYKLIILEVLYSLAKKE 1071
Qy 422 KQLRLDTDQYVGNATYF-----ENGRVVDTCGNSWSEVLQIQETARTIIFNG-KD 473
Db 1072 ESLSLICTSLMETFRKTLNSYKGBDEKINTYLTSDNTQELLK--LRVSLNIDNGYVD 1129
Qy 474 LNIIVERIAAVNPS-----DPLETTKPD----- 496
Db 1130 PDLRLKQSRQAQSSSTQESSGSANGKATATCAGSLETSTSTNVPTVFAGGPRDSHNSVGSIG 1189
Qy 497 ---MTLKEALKIAFGNPNGL-----OYQOKDITEFDFFDOOTSQ 537
Db 1190 FPNMNIQGRRTSGFS--PRVVKKSSLSKEIDPKTFYEYEPKEGKSFDNDNDQDTNIG 1248
Qy 538 IKN-QLAEL-----ATNIYT-----VLDKIKLNKKNILIRDRKFRHYDRN 577
Db 1249 SFNLHLFDMYGISISSSTNSISSSDBLESKEQEQLQDLLETEREDSAEILDARF---RN 1305

```

```
QY 578 NIAVGADSVVKEAHREVINSTEGLLNIDKIRKILSGVIVIEDTEGLKEVINDRYD 637
Db 1306 K-----EDDKVTKD-----ISNDKKNYLPHEKNILKAKEG--KETRDVREBESEDFDLG 1354
QY 638 MNISLSLQDGTFTDFKYNKDL-PLYISN--PNYKVNVTAVT 678
Db 1355 MSLNKKIKREAR-HVD--KVDDSDVDFPKSSAFQSPPIRAYGST 1395

RESULT 12
RBPI_PLAVB STANDARD; PRT; 2869 AA.
AC Q00758;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE Reticulocyte binding protein 1 precursor.
GN RBPI.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
merozoites."
RL Cell 69:1213-1226(1992).
CC -!- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to
human reticulocyte cells.
CC -!- SUBUNIT: Homodimer (Potential).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
Db 2109 EMNTYNSIT-----QLEGIIVVSAGESKEDIKLEERSNEWRNISEKIS 2152
QY 295 TSTRHTSNSNSTVAIDHSLSLAGERTWAETMGLNTADPARLNANIRYV-----NTGT 347
Db 2153 TIDSKVIENNST---IDELYKL-GKNCQAHWISL-ISTANMKTSKLLIMINKENTSEK 2207
QY 348 APIYVNLPTSLVLGKNQTLAI---KADENQLSQILAPNNYYPKSLA---PIALNAOK 401
Db 2208 CVDY-IFDNSSSTDGVVETLKGFGYKLTFFSASEIVQVADTY-SVNFAPKHEKESLNAR 2265
QY 402 DASSTPITMYNQFLEL-----EKTQRLRLDQVYGNIAFYNFENGRVRVDT 449
Db 2266 DIKKELYLPHONSDISIVEGGVQNMLALYDLKNEEKREWDLYRNISETKLQMEHSTDV 2325
QY 450 GSNWSEVLPIQIETTARIIFNGKDLNLVRRATAVNPSPDLETTPKPDWTLKALKI---- 505
Db 2326 FKPMIELHKGMNET-----NNKSLEKEKKLQSVN--DHMSHMEABM-IRNGLKYTPES 2376
QY 506 AFGFNEPNNGNLQYQKDIITEFDNF--DQQTSGNIKNQLAEL-NATNIYTVLIDKIKLNK 562
Db 2377 VQNNINNIYSVIEAEVKTLEEDRDYGDNYQIVVEHKKQFSILDRN--ALMDDIEIFKK 2434
QY 563 MNILIRDKRFHY---DRNNIAGADESVVEAHREVINST--EGLLLNI---DKDIRKI 614
Db 2435 EN-----NYNLMVNTTETHRVNDYIEKTNKLQVQAKTYEQILENIKQNDMLQNI 2486
QY 615 LSGYVIEIEDTEGLKE---VINDRYDM-----LNISSLRDGKGTDFDKKYNKDLPL 663
Db 2487 FLKKVSIIEYFENVKVKESILNDLYQERLLAKIGELHDEIKENVTEILSSVEIQKMEM 2546
QY 664 YISN-----PNYKVNVTAVTKNTIINPS-----ENGDTSTNGIKKILIFSKKG 707
Db 2547 MSKNLLEKSKWMNY-TSYELEREANEINRDAKQIKDDDTILNSVLEAAI-QKRG 2600

RESULT 13
BXEN_CLOBU STANDARD; PRT; 1162 AA.
AC Q06366;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Botulinum neurotoxin type E, nontoxic component.
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BL6340.
RA MEDLINE=93360835; PubMed=8355622;
RA Fujii N., Kimura K., Yokosawa N., Oguma K., Yashiki T.,
RA Takeshi K., Ohnaka T., Isogai E., Isogai H.;
RT "Similarity in nucleotide sequence of the gene encoding nontoxic
component of botulinum toxin produced by toxigenic Clostridium
butyricum strain BL6340 and Clostridium botulinum type E strain
Mashike."
RL Microbiol. Immunol. 37:395-398(1993).
CC -!- FUNCTION: THE NONTOXIC COMPONENT IS NECESSARY TO MAINTAIN
TOXICITY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```



```
DR InterPro; IPR000395; Peptidase_M27.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR ProDom; PD001963; Bontoxilysin; 1.
KW Neurotoxin.
SQ SEQUENCE 1162 AA; 136829 MW; C86E9BE932DA78E4 CRC64;

Query Match 4.4%; Score 159.5; DB 1; Length 1162;
Best Local Similarity 21.5%; Pred. No. 1.3;
Matches 173; Conservative 111; Mismatches 288; Indels 231; Gaps 43;

QY 2 VKQENRLNESSSOGLGYFSDLNFOAPMVVTSITGDLISIPSS-----ELENI 53
Db 372 INQNTVLMKSNFYGDGLG---NDNFVNTIIPYNLYNHSINYSKLDNNVIEEIKI 428
QY 54 PSENO-----YFQSAIWSGFI---KVKSDYEYFATSAD--NHTVMWDDQEVINKASNSN 104
Db 429 PPINDEDIYPRKNA--DTFIPVNIITAKAINTTTPLPVNVLOAQMIDSDNLSDFL 486
QY 105 KIRLEKGL-----YQIKIOYORENPEK-----GLDFKLYTWDQNKKEVIS- 147
Db 487 KVISKGLSVYFLANTMDYLFYKDYKPIDIDTKKYKWLKAFRNSLDITQETISNQ 546
QY 148 -SDN-LQLPELQ-----KSSNRKRKST-SAGP-----TVPDRNDGIPDSLEVEG 191
Db 547 FGDTKIIPWIGRALNLTNTNSFVEEFKULGPFLINKKENITIPKIDRIPSSM----- 602
QY 192 YIVDVKNKETFSPWISNIEHKG--LTK-YKSSPEKSTASDPYSDFEKVTGRIDKQVS 248
Db 603 LNFSPKD-----USENLFNYCKNNFYKRIYFNLDQWMT--QYYSQYFDLICKASKVL 656
QY 249 PEARHPLVAAPIVHVDMENIILSKNEDQSTQNTD--SETRTISKTSTSRHTSNSNST 307
Db 657 --AQELKIK-----LIQQLRYLMENISSTNLILNLTNTLRDISNQSQ 703
QY 308 VAIDHLSLAGETWAEWTGLTADTARLANI--RYNTGTAPYVNVLPPTSILVGLKQK 365
Db 704 IAINNIDKF-----FNNAACVFEENIYKPT-SFMEQCCKNKST-----KEF 747
QY 366 TLATIKADENQLSQILAPNPNYPSKNLAPIALNAQKASSTPTVMYNOFLELEKTKOLR 425
Db 748 ILKCTINETERKSHLMQNSP-----SNLDFD-----FLDIQNMKNL- 784
QY 426 LTDQVYGNIAYNFENGVRVDTSNNGSEVLPOIQTETARII--FNGKDLNLVERRTAA 483
Db 785 -----FNLYTELLIKEQTSPEYLSLYAFQEQDNNVIGTSGKN-TLVE----- 826
QY 484 VNPSDELETTKPDWTLKEALKIAGFNEBNGNLQYQKDIETFDNFQQTQSONIK--NQ 541
Db 827 -YPKD-----IGLVYGIN--NNAIHLTG-----ANQNIKFTND 856
QY 542 LAELNATN---IYTVLDKIKLAKNKNILIRDK-----RFHYDRNIAV-----GADES 586
Db 857 YFENGLTNFYSIYFMRNLNQNTIKSKLIGSKEDNCGWEIYFENGLVFNIDSNGNEKN 916
QY 587 VV-----KEAREV--INSSTEGLLNID-----KQIRKILSYIVIEID----- 624
Db 917 IYLSNISKSWHYVISINRLKQQLIFIDNLTIVANEDIKEILNYSDDIISLSDNNNV 976
QY 625 -TEGL-----REVINDRYDMLNISLSRODGKTFIDFKKYNDKPLIYISNPNYKVV 674
Db 977 YIEGLSVLNTKINSNELLTDYFSDLNNSVIRNFDEIL--QYNHTYELF-----NYVPE 1029
QY 675 YAVTK--ENTIINPENGDTSTN 695
Db 1030 IAINKIEQNNIYLSNNNSLN 1052

RESULT 14
ID SIR4_YEAST STANDARD; PRT; 1358 AA.
AC P11978;
DT 01-OCT-1989 (Rel. 12, Created)
```



```
CC or send an email to license@sib.ch).
CC -----
DR EMBL: M14628; AAA3227.1; -.
DR FIR; A2655; A2655.
DR PDB; 1MMA; 03-DEC-97.
DR PDB; 1MMD; 17-AUG-96.
DR PDB; 1MMG; 03-DEC-97.
DR PDB; 1MMN; 03-DEC-97.
DR PDB; 1MND; 17-AUG-96.
DR PDB; 1MNE; 17-AUG-96.
DR PDB; 1VOM; 23-DEC-96.
DR PDB; 1LVK; 28-JAN-98.
DR PDB; 1DOX; 20-DEC-00.
DR PDB; 1DOY; 20-DEC-00.
DR PDB; 1DOZ; 20-DEC-00.
DR PDB; 1DIA; 20-DEC-00.
DR PDB; 1DIB; 20-DEC-00.
DR PDB; 1DIC; 20-DEC-00.
DR PDB; 1FNV; 20-DEC-00.
DR PDB; 1FMW; 20-DEC-00.
DR PDB; 1G8X; 17-JAN-01.
DR PDB; 1JWY; 07-NOV-01.
DR PDB; 1JX2; 07-NOV-01.
DR DictyBase; DDB0002015; mhca.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR008989; Myosin_S1_N.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00663; myosin head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PSS0096; IQ; 1.
DR Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
KW Calmodulin-binding; Methylation; Phosphorylation.
FT DOMAIN 1 761 MYOSIN HEAD-LIKE.
FT DOMAIN 762 791 IQ.
FT DOMAIN 817 2116 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 638 660 ACTIN-BINDING.
FT DOMAIN 738 752 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (DI-) (POTENTIAL).
FT MOD_RES 1823 1823 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 1833 1833 PHOSPHORYLATION (BY MHCK).
FT MOD_RES 2029 2029 PHOSPHORYLATION (BY MHCK).
FT TURN 3 5
FT TURN 7 8
FT TURN 10 15
FT HELIX 25 28
FT STRAND 34 37
FT TURN 41 42
FT STRAND 48 55
FT STRAND 59 63
FT STRAND 69 73
FT HELIX 74 76
FT STRAND 78 79
FT HELIX 83 85
FT TURN 86 87
FT STRAND 90 90
FT HELIX 91 93
FT TURN 99 110
FT TURN 111 113
FT STRAND 116 119
FT TURN 120 121
FT STRAND 122 126
FT HELIX 137 142
FT TURN 143 145
FT HELIX 148 150
FT HELIX 155 169
FT TURN 170 170

Query Match 4.4%; Score 159.5; DB 1; Length 2116;
Best local Similarity 19.3%; Pred. No. 2.8;
Matches 131; Conservative 116; Mismatches 232; Indels 201; Gaps 33;
QY 86 HVTWVDDQEVINKASNSKIRLEKGRLYQIKIQYORENPTEKGLDFKLYWTDSONKKE- 144
DB 1294 HVNEQUEEEKKQKSEKKEKVDLEK-EVSELKDKQIEEFVASKAV-----TEAKNKES 1346
QY 145 -----VSSDNQLPELKQKSSNSRKRSTSGPTVPDRNDGIPDSLEVEGYT 193
DB 1347 ELDEIKHQYADVSSRDKSVQELTKLQAKNEELRNTA-----BEAQGLDRAERS--- 1396
QY 194 VDVKKRTF-LSPWISNIHEKGLTKYKSPKSWSTADPSYDFEKTGRID--KNVSPE 250
DB 1397 ---KKKAEFDLEEAHVKNLEEE---TAKKVAEKAMKKAEE--IDYRSTKSELDDAKNVSE 1448
QY 251 ARHPLVAAPYIVHVDMENIILSKNEPQS-----TQNTDSETRTISKNTSTSRTHTSNS 305
DB 1449 Q-----YVQIKRL---NEELSELRSVLEBADERCNSAIKAKKTAESALESKD 1493
QY 306 STVAIDHSLSLAGERTWAEWTGLNTADTARL-----NANIRYVNTGTAPIYVNLPTSL 359
DB 1494 EIDAANNAKAKAEK--SKELVRVAELESLEDKSGTVNVEFIRKDAEIDDLRLDL- 1550
QY 360 VLGNQTLATIKADENQLSQILAPNNYPSKNLAPIALNAQKSDASSTPTITMNYNQFLELE 419
DB 1551 ----RETESRIKSDCK-----KNTRK-----QFADLE 1574
QY 420 KTKQLRLDTQVYGNIAIYFNENGVRVVDGTSNWSVLPQIQ-ETTARI-IFNGK---DL 474
DB 1575 -----AKVEEAQREVVTID-----RLKKKLESIDIILSTQDLETTSRIKESKKLEQ 1624
QY 475 NLVERRIAAVNPSPLETTKPKDMTLKALKIAFGFNEPNGNLQVQGGKIDTEFDNFDQ- 533
DB 1625 TLAERRAAEEGSSKAAD-----EIRKQVW-----QEVDLRAQLDSE 1663
QY 534 -----TSQNTKNQALAEALNATNYTVLDKIKLNAMNILLIRDKRPHYDRN---NIAVGADE 585
DB 1664 AALNASEKKIKSLVAE-----VDEVKEQLEDEILAKDKLVKAKRALEVELEVRDQ 1714
QY 586 SVVKEAHREVINSSTEGLLNIDKIDKILSGYIVEIE-----DTEGLKEV 631
DB 1715 LEEEDSRSELEDSKRLTTEVE--DIRK---KYDAVEQNTKLDEAKKLTDDVDTLKKQ 1770
QY 632 INDYDMLNISS-----LRQDKTFI-----DFKKYNDKLPYISPNYK 671
```

Db 1771 LEDEKKKLNESERAKRLESENEDEFLAKLDAEVKXRSRAEKORKKYEKDL-----KOTKYK 1826  
Qy 672 VNVYAVTKENTIIINPSENGD 691  
Db 1827 LNDERATKTQTEIGAACLE 1846

Search completed: May 3, 2004, 19:37:07  
Job time : 10.9844 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 33.4955 Seconds  
(without alignments)  
6697.424 Million cell updates/sec

Title: US-09-848-909A-16  
Perfect score: 3648  
Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKKILFSKKGYEIG 711

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3535	96.9	741	2 Q937W2	Q937W2 bacillus an
2	3535	96.9	743	2 Q937W3	Q937W3 bacillus an
3	859.5	23.6	876	2 Q32739	Q32739 clostridium
4	857.5	23.5	876	2 Q9KH41	Q9KH41 clostridium
5	849.5	23.3	879	2 O06498	O06498 clostridium
6	849	23.3	875	2 Q45221	Q45221 clostridium
7	792	21.7	721	2 O86171	O86171 clostridium
8	705	19.3	775	2 Q844J8	Q844J8 bacillus th
9	296.5	8.1	225	2 Q8KYK2	Q8KYK2 bacillus th
10	218.5	6.0	1881	16 Q8RGK2	Q8RGK2 fusobacteri
11	205.5	5.6	675	16 Q897H6	Q897H6 clostridium
12	203.5	5.6	1301	5 Q8SKS5	Q8SKS5 plasmodium
13	200	5.5	4688	16 Q9PQ08	Q9PQ08 ureaplasma
14	199.5	5.5	2849	5 Q8IHY4	Q8IHY4 plasmodium
15	193.5	5.5	3468	5 Q8II04	Q8II04 plasmodium
16	199	5.5	1404	2 Q8RJN9	Q8RJN9 mycoplasma

17	198.5	5.4	1301	5 Q8IHQ2	Q8IHQ2 plasmodium
18	197	5.4	1302	2 Q49547	Q49547 mycoplasma
19	196.5	5.4	1904	5 Q8IKS2	Q8IKS2 plasmodium
20	196	5.4	3248	5 Q8ICP9	Q8ICP9 plasmodium
21	194	5.3	1859	5 Q8IC27	Q8IC27 plasmodium
22	194	5.3	2792	5 Q8I4R2	Q8I4R2 plasmodium
23	192.5	5.3	2529	16 Q25579	Q25579 helicobacte
24	192.5	5.3	2771	5 Q26216	Q26216 plasmodium
25	192.5	5.3	5767	5 Q8I525	Q8I525 plasmodium
26	192	5.3	1833	5 Q8IAZ5	Q8IAZ5 plasmodium
27	192	5.3	2940	5 Q8IHP9	Q8IHP9 plasmodium
28	192	5.3	9904	16 Q8NWO6	Q8NWO6 staphylococ
29	190.5	5.2	1263	5 Q8IK49	Q8IK49 plasmodium
30	190.5	5.2	1722	5 Q8II18	Q8II18 plasmodium
31	190	5.2	752	16 Q98PL9	Q98PL9 mycoplasma
32	189.5	5.2	2577	5 Q8I531	Q8I531 plasmodium
33	188.5	5.2	4095	2 Q7WZ13	Q7WZ13 haemophilus
34	188	5.2	1425	5 Q8IK08	Q8IK08 plasmodium
35	187	5.1	2269	5 Q26223	Q26223 plasmodium
36	187	5.1	6713	16 Q99U54	Q99U54 staphylococ
37	187	5.1	6713	16 Q931R6	Q931R6 staphylococ
38	186.5	5.1	990	5 Q8I5C6	Q8I5C6 plasmodium
39	186	5.1	2747	5 Q9BUX9	Q9BUX9 plasmodium
40	185.5	5.1	2616	5 Q8IIG1	Q8IIG1 plasmodium
41	185	5.1	1936	5 Q8I5A6	Q8I5A6 plasmodium
42	184.5	5.1	2578	5 Q8IJP9	Q8IJP9 plasmodium
43	184	5.0	1151	5 Q8ILU0	Q8ILU0 plasmodium
44	184	5.0	2072	5 Q8II32	Q8II32 plasmodium
45	184	5.0	2752	5 Q9BJY0	Q9BJY0 plasmodium

#### ALIGNMENTS

#### RESULT 1

Q937W2 ID Q937W2 PRELIMINARY; PRT; 741 AA.  
 AC Q937W2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Rag protein (Fragment).  
 GN PAG.  
 OS Bacillus anthracis.  
 OG Plasmid pXOI.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1392;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Ferrara;  
 RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,  
 RA Fasanella A., Francia M., Ciuchini F.,  
 RT "Sequence analysis of the genes encoding for the major virulence  
 factors of bacillus anthracis vaccine strain 'Carbosap'.  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ413937; CAC93935.1; -  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
 DR GO; GO:0015070; F:toxin activity; IEA.  
 DR GO; GO:0009405; P:pathogenesis; IEA.  
 DR InterPro; IPR003896; Anthrax toxins.  
 DR Pfam; PF03495; Binary toxR; I.  
 DR PRINTS; PR01391; BINARYTOXINE.  
 KW Plasmid.  
 FT NON\_TER 1  
 FT NON\_TER 741  
 SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 96.9%; Score 3535; DB 2; Length 741;  
 Best Local Similarity 96.0%; Pred. No. 2.8e-160;  
 Matches 693; Conservative 1; Mismatches 4; Indels 24; Gaps 1;

Cy 1 EVKQENLLNESSSQGLLGYFSDLPAPMVVTSSTGDLSPSSLENIPISENQYF 60

```
Db 20 EVKQENRLNSESSESSQGLLYYFSDLNFAQPMVVTSTTGLSLPSSLENIIPSENQYF 79
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 80 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 139
QY 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPLKQKSSNSRKRSTAGPTVPDRDN 180
Db 140 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPLKQKSSNSRKRSTAGPTVPDRDN 199
QY 181 DGIPLDSLEVEGYVVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSWASTASDPYDFEKT 240
Db 200 DGIPLDSLEVEGYVVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSWASTASDPYDFEKT 259
QY 241 GRIDKNVSEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 260 GRIDKNVSEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 319
QY 301 -----SNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 336
Db 320 SEVHGNAEYHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 379
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 396
Db 380 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 439
QY 397 LNAQKDASPTITMNYNOFLEKTKQLRLDQVYGNATYVNFENGRVVDVDTGNSWSEV 456
Db 440 LNAQDDFSPTITMNYNOFLEKTKQLRLDQVYGNATYVNFENGRVVDVDTGNSWSEV 499
QY 457 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 516
Db 500 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 559
QY 517 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 576
Db 560 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 619
QY 577 NNAIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 636
Db 620 NNAIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 679
QY 637 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNVYAVTKENTIINPSENGDSTNG 696
Db 680 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNVYAVTKENTIINPSENGDSTNG 739
QY 697 IK 698
Db 740 IK 741
```

## RESULT 2

```
Q937W3 ID Q937W3 PRELIMINARY; PRT; 743 AA.
AC Q937W3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pag protein (Fragment).
GN PAG
OS Bacillus anthracis.
OC Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Carbosap;
RA Adone R.; Pasquali P., La Rosa G.; Marianelli C.; Muscillo M.,
RA Paganella A.; Francia M.; Ciuchini F.;
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbosap.'";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL: AJ413936; CAC93934.1; -.
DR GO: GO:0005576; C:extracellular; IEA.
DR GO: GO:0046821; C:extrachromosomal DNA; IEA.
DR GO: GO:0015070; F:toxin activity; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR003896; Anthrax toxinB.
DR Pfam: PF03495; Binary toxB; I.
DR PRINTS: PR01391; BINARYTOXINB.
KW Plasmid.
KW NON_TER.
FT NON_TER 743 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5D85A6 CRC64;

Query Match 96.9%; Score 3535; DB 2; Length 743;
Best Local Similarity 96.0%; Pred. No. 2.8e-160;
Matches 693; Conservative 1; Mismatches 4; Indels 24; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFAQPMVVTSTTGLSLPSSLENIIPSENQYF 60
Db 22 EVKQENRLNSESSESSQGLLYYFSDLNFAQPMVVTSTTGLSLPSSLENIIPSENQYF 81
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 82 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 141
QY 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPLKQKSSNSRKRSTAGPTVPDRDN 180
Db 142 QRENTEKGLDFKLYWTDSONKEVIVSSDNLQPLKQKSSNSRKRSTAGPTVPDRDN 201
QY 181 DGIPLDSLEVEGYVVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSWASTASDPYDFEKT 240
Db 202 DGIPLDSLEVEGYVVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSWASTASDPYDFEKT 261
QY 241 GRIDKNVSEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300
Db 262 GRIDKNVSEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 321
QY 301 -----SNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 336
Db 322 SEVHGNAEYHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 381
QY 337 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 396
Db 382 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 441
QY 397 LNAQKDASPTITMNYNOFLEKTKQLRLDQVYGNATYVNFENGRVVDVDTGNSWSEV 456
Db 442 LNAQDDFSPTITMNYNOFLEKTKQLRLDQVYGNATYVNFENGRVVDVDTGNSWSEV 501
QY 457 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 516
Db 502 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPNGNL 561
QY 517 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 576
Db 562 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRFHYDR 621
QY 577 NNAIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 636
Db 622 NNAIAGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 681
QY 637 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNVYAVTKENTIINPSENGDSTNG 696
Db 682 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKNVYAVTKENTIINPSENGDSTNG 741
QY 697 IK 698
Db 742 IK 743

RESULT 3
O32739 PRELIMINARY; PRT; 876 AA.
ID O32739
```

AC O32739;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE ADP-ribosyltransferase.  
GN CDTB.  
OS Clostridium difficile.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1496;  
RN [1]\_SEQUENCE FROM N.A.  
RP STRAIN=CD196;  
RC MEDLINE=97230316; PubMed=9119480;  
RA Perille S., Gilbert M., Bourlioux P., Corthier G., Popoff M.R.;  
RT "Production of a complete binary toxin (actin-specific ADP-  
RT ribosyltransferase) by Clostridium difficile CD196.";  
RL Infect. Immun. 65:1402-1407(1997).  
DR EMBL; L76081; AAB67305.1; -;  
DR HSP; P13423; IACC.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax toxinB.  
DR Pfam; PF03495; Binary toxB; I.  
DR PRINTS; PR01391; BINARYTOXINB.  
DR Transferase.  
SQ SEQUENCE 876 AA; 98797 MW; 255062D45C2B3B CRC64;  
Query Match 23.6%; Score 859.5; DB 2; Length 876;  
Best Local Similarity 32.1%; Pred. No. 5.7e-33;  
Matches 251; Conservative 131; Mismatches 277; Indels 123; Gaps 31;  
QY 4 QENRLNSESSESSQGLGYFSDLNFAQPMVTSSTGDLSPSELEN-IPSENQYFQS 62  
DB 40 KKKEIVNEIDLPNGLMGYFSDHEFKDLKMAPDKGNLKEFKKVDKLDKDKSDVKS 99  
QY 63 AIWSGFIVKKSDEYTFATSDNHNVTMVDQEVINKASNKIRLEKGRLYQIKIQYOR 122  
DB 100 IHWTGRIIPSKDGEYTLSTDRDD-VLMQVNTESI---SNTLKNVMMKKGKVKRIELQD 155  
QY 123 EN--PTEKGLDFKLWTSQNKKEVSSDNIQLPELKOKSSNKRKSTSGAPVTP--- 176  
DB 156 KNLGSDNLSSPNLYW-ELDGMKKIIPENLFLRDY-----SNIEKD-----DPFIPNNNF 205  
QY 177 -----DRDNDGIPDSLEVEGYTVDVKNKTFLSPWLSNIHEKKGLTKYKS 222  
DB 206 FDPKLSMDEDELDTDNDNIPDSYERNGYTI---KDLIAVKWEDSPAE-QGYKKYVSN 260  
QY 223 PEKWTASDPYDFEKVTGRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEQDST--- 279  
DB 261 YLESNTAGDPYTDYEKASGSPDKAITEARDPLVAAPYIVGVGMEKLIISTNEHASTDQG 320  
QY 280 -----QNTDSETRTISKN-----TSTSRHTSNSSSTVAIDHSLSLAGERTWA 323  
DB 321 KTVSRATNKSSTESAGVNVGVQNGFTANVTNYSHTDNTSTAVQDS---NGE-SWN 376  
QY 324 ETWGLNTADTARLANIRYVGTAPIYNNLPTTSLVLGKQNTLATIKADENQSLIAP 383  
DB 377 TGLSINKGESAVINAVRYNGTAPMYKVPTTNVL-DGDTLSTIKAEQNGIINNLS 435  
QY 384 NNYPSKNLAPIALNAQKADASTPTIMYNGFLEKTKYQLRDTDDQYVGNIAFYENG 443  
DB 436 GDTYPPKGLSLPALNMDQFSRLIPINYDQLKLDAGKQIKLETQVSGNFGTKN-SSG 494  
QY 444 RVRVDTGSNWSVLPOIETTARIIFNGKDLNVLVERIAAVNPSPDLETTKPDMLKEAL 503  
DB 495 QI-VTEGNSWSDYISQDISASIIIDTEN-ESVRRVTAQNQDPEKDT-PELTIGEAI 551  
QY 504 KTAFGNPNGLQYQGGKITE--PDFNPDQTSQNTKQKQAEIATNIIYTVLDKIKLNA 561  
DB 552 EKAFGATKDGGLYFNDFIDBSCVELIFDNTANKIKDLSKTLSDKKIYNV-----KLER 607

QY 562 KXNILIRDKRPHY---DRNNAVGADSVVKEAHREVINSSTEGL-----LINIRK 609  
DB 608 GWNILIKTYTTFNFDYNNYP--STWSNVNTNQDGLQGSANKLNGETKIKIPKSELKP 665  
QY 610 DIRKILSGY-----IVEI-----EDTEGL-----KEVINDRYMLNIS-SL 644  
DB 666 YKRYVFSGSKDPLTSNSIIVIRKAKEBKTDYLVPEQGYTKFSYFETTEKSSNIEITL 725  
QY 645 RQDGKTFIDPKYND--KLPLVISPNYKV-----NVYAVTKNTIINPSENGDTSTN 695  
DB 726 IGSYTYLDNLSTELNSTFEILDPEVKIPTDQEIIMDAHKIYFADLNFPS-TGNTIYN 784  
QY 696 GI 697  
DB 785 GM 786  
RESULT 4  
Q9KH41  
ID Q9KH41 PRELIMINARY; PRT; 876 AA.  
AC Q9KH41;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE CdtB.  
GN CdtB.  
OS Clostridium difficile.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1496;  
RN [1]\_SEQUENCE FROM N.A.  
RP STRAIN=CUG 20309;  
RC Chang S.Y., Song K.P.;  
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain  
RT CUG 20309.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP21719; AAF81761.1; -;  
DR HSP; P13423; IACC.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax toxinB.  
DR Pfam; PF03495; Binary toxB; I.  
DR PRINTS; PR01391; BINARYTOXINB.  
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;  
Query Match 23.5%; Score 857.5; DB 2; Length 876;  
Best Local Similarity 32.0%; Pred. No. 7.1e-33;  
Matches 250; Conservative 132; Mismatches 277; Indels 123; Gaps 31;  
QY 4 QENRLNSESSESSQGLGYFSDLNFAQPMVTSSTGDLSPSELEN-IPSENQYFQS 62  
DB 40 KKKEIVNEIDLPNGLMGYFSDHEFKDLKMAPDKGNLKEFKKVDKLDKDKSDVKS 99  
QY 63 AIWSGFIVKKSDEYTFATSDNHNVTMVDQEVINKASNKIRLEKGRLYQIKIQYOR 122  
DB 100 IHWTGRIIPSKDGEYTLSTDRDD-VLMQVNTESI---SNTLKNVMMKKGKVKRIELQD 155  
QY 123 EN--PTEKGLDFKLWTSQNKKEVSSDNIQLPELKOKSSNKRKSTSGAPVTP--- 176  
DB 156 KNLGSDNLSSPNLYW-ELDGMKKIIPENLFLRDY-----SNIEKD-----DPFIPNNNF 205  
QY 177 -----DRDNDGIPDSLEVEGYTVDVKNKTFLSPWLSNIHEKKGLTKYKS 222  
DB 206 FDPKLSMDEDELDTDNDNIPDSYERNGYTI---KDLIAVKWEDSPAE-QGYKKYVSN 260  
QY 223 PEKWTASDPYDFEKVTGRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEQDST--- 279  
DB 261 YLESNTAGDPYTDYEKASGSPDKAITEARDPLVAAPYIVGVGMEKLIISTNEHASTDQG 320  
QY 280 -----QNTDSETRTISKN-----TSTSRHTSNSSSTVAIDHSLSLAGERTWA 323







```
QY 286 TRTSKNTSTSRHTNSNSSTVAIDHSLSLAG-----BRTWAET----- 325
Db 287 TKSMKSTSHSST--NINTVGAEVSGSLQAGGIPFVPSMSASANYSHTWNTSTVDVT 343
QY 326 -----NGLTADARLANAIRYNTGTAPIYVNLPTSLVLKGNQTLATIKADENQL 377
Db 344 TBSFQGSQSGSINTGESAYINPIRYNTGTAPIYVNLPTSLVLKGNQTLATIKADENQL 402
QY 378 SOILAPNNYPPSKNLAPIALNAQKQDASSTPTIMYNNQFLELEKTKQLRLDQVYGNAT 437
Db 403 GOVNLPGGYPIIGEPMALNTMDQSSSLIPINYNQLKSIDNGTVMVLSSTQFTGNFAX 462
QY 438 YNFENGRVRVDGNSWSEVLPQIQTETARI--FNGKOLNLVERIAAVNSDPLETTKP 495
Db 463 YN-SGNLVT-D--GNWGPYLGITKSTASLTLSFGSGTQVA--VWAPNFSDDPKT-P 516
QY 496 DMTLKEALKIARFENPNGLQYQKDI--EFDFNFQDQTSQNIKNQLAELNATNIYT 552
Db 517 KLTLEQALVKAFALEKKNKGFYHGLEISKNEKIQVFLDSNTNDFENQLKNTADKDIWH 576
QY 553 VLDIKIKNAKMILIRDKGFHYDRNNAIV 581
Db 577 CI--IKRN--MNILVKVITKENISSINI 601

RESULT 8
Q844J8 PRELIMINARY; PRT; 775 AA.
AC Q844J8;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE ViPlac.
GN ViPlac.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Pang Y.;
RT "Cloning of vipA(C) and vip2A(C) from Bacillus thuringiensis HD201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY245547; A086514.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR001896; Anthrax toxinB.
DR Pfam; PF03455; B:toxinB; I.
DR PRINTS; PR0139; BINARYTOXINB.
DR SQR SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;

Query Match 19.3%; Score 705; DB 2; Length 775;
Best Local Similarity 28.9%; Pred. No. 1.1e-25;
Matches 220; Conservative 129; Mismatches 285; Indels 128; Gaps 29;

QY 4 QENRLNESSSQGLCYFSDLNFAQPMVYTSSTGSLSPSSELEN--IPSENQYFQ 61
Db 42 QKNQ---QKMDRKGGLGYFFKQDF-SNLTMAPTRDNTLYDQQTANKLDRKQYQ 97
QY 62 SAWSGFYKVKSDYTFATASDNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIQY 121
Db 98 SIRWIGLQSKGKDETFNLSEDEQAIIDGKLIISNKGKQVHVLEKLVPIKIEVQ 157
QY 122 RENPTKGLD-----FKYWTDSQNKKEVISEDNIQLBELKQKSN----- 162
Db 158 SD--TKFNIDSKTFKEFKLFKIDSONQSQVKRDELNPFENFKGSEREFLAKASKNTFFM 215
QY 163 SRKRSSTAGTVPDRNDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 216 QKMKRDIDED--TDTGDSIPDLWEENGYT--IQNK--VAVK--DDKFAQQYGVKYLSS 267
QY 223 PEKWTASDYSDFEKVTGTIDKKNVSPEARHPVAAAYPIVHVDMENILSKNEDOSTQNT 282
```

```
Db 268 PYQARTVGDPTDWEKAAGDIPKSNAAATRNPLVAAFPSPINVDMMRWILSKDNLNLS----- 323
QY 283 DSETRITSKNTSTSRHTNSNSSTVAIDHSLSLAG-----BRTWAET----- 327
Db 324 -----NSAEHSNSSYTYANSEGASIEAGFGPKGFGVGSANTQHTETVGSWDWG 372
QY 328 -----LNTADARLANAIRYNTGTAPIYVNLPTSLVLKGNQTLATIKADENQLSQ 379
Db 373 NSKSNTQEFNSAGVILNANVHYNNVGTGGIDYDQFTTSPIL-QDSTIATITAKSNATAL 431
QY 380 ILAPNNYPPSKNLAPIALNAQKQDASSTPTIMYNNQFLELEKTKQLRLDQVYGNATYN 439
Db 432 SIFSGRYPASK-EGSLKTMDDFNHPTLNKPOLDAVLNNEVIKINTDQDGRYGIIG 490
QY 440 FENGRVRVDGNSWSEVLPQIQTETARI--FNGKOLNLVERIAAVNSDPLETTKPDMTL 499
Db 491 VDG---KAEIGDRWSPTIIDEIKGRTASIIIDPADGKALETRIAADKYKNPEDXT-PSLTI 546
QY 500 KEALKIARF--FNEPGLNQYQKDI--EFDFNFQDQTSQNIKNQLAELNATNIYT 545
Db 547 KEGLKIAYPESEDKDGILFYEKNDGKVKTKQUSEENIMPYLDEDTSKFERQUSDG 606
QY 546 NATNIYTVLDDIKIKNAKMILIRDKGFHYDRNNAIVGADESVVYKEAHREVINSSTGLL- 604
Db 607 SAKGLY---DIKLTQKNITIR-----LATVTLGDDQFSAYPWEATWSDKFGNLR 655
QY 605 -----LNIDKDIRK-----ILSGVIVE--IEDTEGLKEVI--NDRYDMLNLS 643
Db 656 LGSALTPQESKYTIPKDKVKNYDILITGYIKHDFTTDNESIGIVAFTKKDNFEMNMGT 715
QY 644 LRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKEINTIIN 685
Db 716 SIFSQNSGGFKEFTIKTQ-NISS-DYILDSIQLMKRNDVN 755

RESULT 9
Q8KYK2 PRELIMINARY; PRT; 225 AA.
AC Q8KYK2;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Protective antigen-related protein, (pXOI-III).
GN BXA0163.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2012;
RX MEDLINE=22061436; PubMed=12004073;
RA Read T.D., Salzberg S.L., Pop M., Shumway M., Umayam L., Jiang L.,
RA Holtzapfel E., Busch J.D., Smith K.L., Schupp J.M., Solomon D.,
RA Keim P., Fraser C.M.;
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RT Bacillus anthracis.";
RL Science 296:2028-2033(2002).
DR EMBL; AE011190; AAM26108.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 225 AA; 25402 MW; 2E121BBE54295F9C8 CRC64;

Query Match 8.1%; Score 296.5; DB 2; Length 225;
Best Local Similarity 37.2%; Pred. No. 5.9e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;

QY 542 LAENANNTVLDKIKNAKMILIRDKGFHYDRNNAIVGADESVVYKEAHREVINSST 601
Db 1 MESLGINNIYNALDRIKNAKMILIRDP-YHYDNNKGNVGVDDSYLKNAKYLWSSD 59
QY 602 GLLNIDKDIRKILSGYVIEPTE-----GLKEVINDRYDKMLNTSSLRQDG 648
```

Db 60 GVSINLDEVDNQALSGYMLQIKKPSNHLTNSPVTITLAGKDSGVGELYRVL-----DG 113

QY 649 KTFIDFKYNDKPLDYISNPNYKVNVAVTYKEN-TIINPSENGDTSTNGIKKILISKKG 707

Db 114 TGFIDFNKFDENWRLV-DPGDDVYVAVYTKEDFNATVDRDENGIA-NKDKNTLVLSGKI 171

QY 708 YEI 710

Db 172 KEI 174

RESULT 10

Q8RGK2 PRELIMINARY; PRT; 1881 AA.

AC Q8RGK2

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hemolysin.

GN FN0291.

OS Fusobacterium nucleatum (subsp. nucleatum).

OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;

OC Fusobacterium.

OX NCBI\_TaxID=76856;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 25586;

RX MEDLINE=2186394; PubMed=11883109;

RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,

RA Bhattacharya A., Bartnan A., Gardner W., Grechkin G., Zhu L.,

RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,

RA Larsen M., D'Souza M., Walunas T., Fusch G., Haselkorn R.,

RA Forstein M., Kyripides N., Overbeek R.;

RT "Genome sequence and analysis of the oral bacterium Fusobacterium

RT nucleatum strain ATCC 25586."

RL J. Bacteriol. 184:2005-2018 (2002).

DR EMU; AE010541; AAL94497.1; -

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR008638; Haemagg.act.

DR InterPro; IPR005058; Peptidase\_S26.

DR InterPro; IPR001452; SH3.

DR Pfam; PF05860; Haemagg.act; 1.

DR Pfam; PF00018; SH3; 1.

DR PROSITE; PS00761; SPASE\_I\_3; 1.

KW Complete proteome.

SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;

Query Match 6.0%; Score 218.5; DB 16; Length 1881;

Best Local Similarity 21.4%; Pred. No. 0.046;

Matches 175; Conservative 130; Mismatches 306; Indels 205; Gaps 37;

QY 30 QAPMVVTS-----STTG-----DLSIPSELE---NIPSENOY-FQSAIWSGFI 69

Db 246 QAPVATKGDVVVISKGVYLTQAKRDIKISSTETETIGSKLLAENAINIKSGKTSNG 305

QY 70 KVKSDVEYTFATSADNHVTKMVDQEVN-----KASNSKIRLEKGR 112

Db 306 QIRANNITINGVDSSNLIITNKDITISGLKNSGVNSNLSNKNVVEE-K 364

QY 113 LYQIKI-----QYQRENPTKGLDFKLYWT--DSQNKKEVISSDNLQPLKQKSS 161

Db 365 LSSTKITNLGLSAKEIEKTNIFNSGKLFKNITAKDFKNQGEV-SSENLTITNLENSNK 423

QY 162 -----NSRKKRSTAGTVPDRNDG----- 182

Db 424 INVKENINSIVNKTNAEITSKNLTNNLDRNITIIINVSSGVIANNGKLLVGNITN 483

QY 183 ---IPDSLEVEGYTVDVKNKRTF.LSPWISNTHEKGLTKYKSSPKKSTASDPYSDEKV 239

Db 484 SONTNTATVQGTLDIKNKNSGKILSDNLAKDFSSGNISAKVITTTQELINSCEII 543

QY 240 TGRIDKNVSPEARPLVAAYPIVHVDMEN-----IILSKNEODSTONTDSETRTSKNT 293

Db 544 SNNLSSNNINNSKIFVNGNLKISNNLANSVGISGLELNTSIENTGNITIKNKLTSQNL 603

QY 294 STSRTHTSNSSTVAIDHSLSLAGERTWAETMGINTADTARLANANTRYVNTGTAPIYV 353

Db 604 N-NKNTANVANAGFLDVHKNISVVG-NIKAITMTKNLND-----NSG-----NI 645

QY 354 LPTTSLVLGKQTLATIKADENQLSOLAPNYPYPSKNLAPIALN-----AQKDASS 405

Db 646 L-TNSLTAEINIKGSIITA-XNISQNLVNSGVSISDNIT-VADNITNINIFANEKISA 702

QY 406 TPITVYNYQFLEKTKQLRLDTQVYGNIAFYAFENGVRVDVTGSNNSEV-----L 457

Db 703 DKIS-NSNK-LVAKNTEITKLTND---GNIVVK--ENLKAKDITNSNTIKVGENLNTDKL 755

QY 458 POIQETTAIIIFNGKDLNLVERRIAANVPSPLETTKPDMLTKALKIATAGFNEPENGLO 517

Db 756 QNSKTLIAKNINIEKSLNNINKKITSLANNTSNTDKNNNGIIQAIK-----NIN 805

QY 518 YQKDIITBPDFNFDOOTSQNIKNQLAELNATNIYTVLD---KIKLNKONILIRDKRPHY 574

Db 806 IKTSNDLKLDGKYTANDSLINAKSLENGN-----LENDGKIKFNLGNLNNKISS 860

QY 575 DRNITAV-----GADESIVVKEAHREVINS--TEGLLL-----NI-DKDIR 612

Db 861 SNLNTANEISNNGVNSIGSEANLITANSUKNEGILLFGEGIENKLTGTNTNTGVI 920

QY 613 KILSGVIVEIETEGLEKVINDRYDMLNISILRODGKTF-----IDFKK--YNDKLPY 664

Db 921 SSLGKLTAKDVNDKHIISDNDLITDVSNTKGLLYSTNNKVKDFKFNLDKAEIY 980

QY 665 ISNPYKVNVAVTAKENTIIINPSENGDSTNGIKKI 700

Db 981 SSG-----DITIN-SENG-TFTNRVGD 1001

RESULT 11

Q897H6 PRELIMINARY; PRT; 675 AA.

AC Q897H6

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Membrane-associated protein.

GN CTC00759.

OS Clostridium tetani.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1513;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=Massachusetts / E88;

RX MEDLINE=22457253; PubMed=12552129;

RA Brueggemann H., Baumer S., Fricke W.F., Wier A., Liesegang H.,

RA Decker I., Herzberg C., Martinez-Arias R., Merki R., Henne A.,

RA Gottschalk G.;

RT "The genome sequence of Clostridium tetani, the causative agent of

RT tetanus disease".

RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321 (2003).

DR EMBL; AE015938; AAC03560.1; -.

KW Complete proteome.

SQ SEQUENCE 675 AA; 77585 MW; 5572933FDE209BFB CRC64;

Query Match 5.6%; Score 205.5; DB 16; Length 675;

Best Local Similarity 20.5%; Pred. No. 0.051;

Matches 144; Conservative 106; Mismatches 217; Indels 235; Gaps 34;

QY 42 DLSIPSELENTPSE-NOYFQSAIWSGFIKVKK--SDEYTFATSDADNHVTKMVDQEVIN 98

Db 52 DINIESK--ENTEGKLCIYISKEYEHHIKISSGVSNKYTFIAPVKNQV-----KEMKI 103

QY 99 KASNSNKILEKGRLYQIKIQYQRENPTKGLDFKLYWTDSONKKEVISSDNLQPLKQ 158



RP	SEQUENCE FROM N.A.	
RC	STRAIN=Serovar 3; PubMed=11048724;	
EX	MEDLINE=20500219; PubMed=11048724;	
RA	Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,	
RA	Cassell G.H.;	
RT	"The complete sequence of the mucosal pathogen Ureaplasma	
RT	urealyticum";	
RL	Nature 407:757-762(2000).	
RL	EMBL; AE002145; AAF30894.1; --	
KW	Hypothetical protein; Complete proteome.	
SW	SEQUENCE 4688 AA; 534880 MW; B53ABFAFFEE1997E CRC64;	
QY	Query Match 5.5%; Score 200; DB 16; Length 4688;	
DB	Best Local Similarity 20.3%; Pred. No. 1.1;	
DB	Matches 166; Conservative 127; Mismatches 313; Indels 210; Gaps 37;	
QY	19 LLYYFSDNLQAPWVTSTTGSLSPSSLENIIPSENQYFQSAIWSGRIKVKSDYET 78	
DB	3699 LVDVYLD-NIHQIDETRKIPKHNH-SKEIENPGVTMISKGNKSPDTTANPEFK 3756	
QY	79 FATSADNHTVMWDDQEVINKASNKIRLEKGLYGIKIYQENRTEKGLDFKLYWD 138	
DB	3757 IETQ-----DNDVLANIDATVFKDEHNKIKQIVRIKEN-----ND 3795	
QY	139 SONKEVSSNLQIPKQKSSN-----SRKRSTSGPTVPDRD----- 179	
DB	3796 WLKGGQI---DNLN-PETKYLENIELSKPLKTHNLNSINDKENSLITETGNPVLKV 3851	
QY	180 ---NDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSWSTASDPYSD 235	
DB	3852 IOTQNDTINDTQIINTVLSGVNSK-YNGROIKVYKDNVNIYESS---LITLQKQND 3907	
QY	236 PEKVTGRIDKNVSPARHPLVAAPVIVHVDVNIILSN-EDQSTQNTDSET-RTISKNT 293	
DB	3908 YQLLSNLNSN-----REYFEKIEINHISNTNPFDEKLKGVSNFTITQTKNT 3957	
QY	294 STSRHTSNSNSTVAIDHLSLAG----- 318	
DB	3958 TVQWDSATVIGTGVNFNFKIKSEDKILENNQOVVAFAPKETIERTNLTWLYTPRK 4017	
QY	319 -----BRTWAEITMGLNT---ADTARLANIRVYNTGTAPIYVNLPTSLVLKQNTLA 368	
DB	4018 DVTDFKEGTWADHLSNVFKEETTYKLVKIQFVNKPTKAKNNINSENIVLDNTNSI 4077	
QY	369 -----TIKADENQLSILAPNNYPSKNLAPIALN-AQKASSTPITMAY--NQFLELE 419	
DB	4078 NSNYEFTKVGDKHLNITSSNNVNTNSQTINFTLSGVKKSWSVGKKIKLSYKSNDSI 4137	
QY	420 KTKQLRLDQVYGNIAIYNFENGK-----VRVDTGSNMSEVLPIQIETTARIIFNGKD 473	
DB	4138 HTEVLESNKTOYVILLNKLNRNTYTLIDVKLIDNNNVSDFFKEGNLTNSFITRTSA 4197	
QY	474 LNLVERRAANVPSDPLETT-----KPDWTLKEA-----LKIAGFNEPENGLO 517	
DB	4198 INVNLIEISNRASTNLKSTIKINLNDPDVLRDKQOATIVYGNKQAMGFITVSGNIK 4257	
QY	518 YQCKDITDFDNFDQOOTSONIK-NQLAELNATNYTLVDKTKLAKAOMILIRDRKRFYD- 575	
DB	4258 YLTALVDLNFN-DKVNIVNIFPNKPSIAEN-----IGDKSNII-----YVND 4304	
QY	576 -----RNNIAGA---DESVVKEAHREVINSSTEGLLNLTDKDIRKILSYGIVIEIDTE 626	
DB	4305 IPKLEINNDIIVNGPINEIIVVGNANQK--NNIDVDLGLQINPKIAHLNR-FIAKFKSTN 4361	
QY	627 GLKEVIND--RYDMLNLSL-RQDGKTRIDF-----KKY----- 657	
DB	4362 -----NDLIETVINGSLSVNDGKTSIRFTLNLKANKLYSLVDVYLVNNSNTIVE 4415	
QY	658 NDKFLYISNPYKVVAVTKENTIIINPSENGDTS 693	
DB	4416 SNKLP-KLNNINYQIK--INKSHITII--SKNGEWS 4445	

RESULT 14	
Q8IHY4	PRELIMINARY; PRT; 2849 AA.
AC	Q8IHY4;
DT	01-MAR-2003 (TREMELrel. 23, Created)
DT	01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT	01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE	Hypothetical protein.
GN	PF11_0392.
OS	Plasmodium falciparum (isolate 3D7).
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX	NCBI_TaxID=36329;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=3D7;
RC	MEDLINE=22255705; PubMed=12368864;
RA	Gardner M.J., Hall N., Fung E., White O., Barriman M., Hyman R.W.,
RA	Carlton J.M., Pain A., Nelson K.S., Bowman S., Paulsen I.T., James K.,
RA	Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA	Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA	Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA	Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA	McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA	Fraser C.M., Barrell B.J.;
RT	"Genome sequence of the human malaria parasite Plasmodium
RT	falciparum";
RL	Nature 419:498-511(2002).
DR	EMBL; AE014841; AAN35975.1; --
KW	Hypothetical protein.
SW	SEQUENCE 2849 AA; 335916 MW; B5515D173D96B813 CRC64;
QY	Query Match 5.5%; Score 199.5; DB 5; Length 2849;
DB	Best Local Similarity 20.6%; Pred. No. 0.63;
DB	Matches 169; Conservative 128; Mismatches 322; Indels 201; Gaps 39;
QY	3 KQENLLNSESQGLLGY-YFSDLNQ--APVVTSTTGDLSP--SSLENIIPSEN 57
DB	1862 KHNVTFFKEENLKSIFKYNNNNNINEKDDIPKIQSSFINEDNIAYNEGNMUKLN 1921
QY	58 Q-YFQSAIWSGFIKVKSDYETFATSADNHTVMWYDD-----QEVINKASNSK---- 105
DB	1922 EDIIQERFHHNIKTVMNKNCSYSDCNIVKINYDELNDSTQTKELNEGKSNCKAEAW 1981
QY	106 -IIEKGRLY-OIKIYQRENTEKGLDFKLYWDSQNKKEVISSDNLQIPKQKSSNS 163
DB	1982 IIDIKNETHYPIKIDKKEKNEDEK--KRYMYNNKNDKNINMKGSN---KSMKNNKNS 2036
QY	164 RKKRSTSGPTVPDRDNDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSP 223
DB	2037 NKMKH-----IPLSVNKKGY-----NKSSINKKYENNIN-----KXNK 2069
QY	224 EKWSTASDPYSDF--EKVTGRI-DKNVSPARHPLVAAPVIVHVDVNIILSKNEDQSTQ 280
DB	2070 DKLNILVNSISKLVOSKIKQELSNKNISK-----DILNFEITKIKKRSKK 2114
QY	281 NTDSETRTSKNTSRTSTNS-----NSSTVA-----IDHLSL 316
DB	2115 ETKNTHNTKNNKNN 316
QY	317 AGE-----RTWAEITMGLNTA-----DTARLANA-----NIRVNTGTAPIYVNLPTSLV 360
DB	2175 APDGTIYVWVNNIDTVMYNYKVFDSKKNINHEVPLANNVPYLN-----NDLLTNVNI 2227
QY	361 LGKNQTLATIKADENQLSQ-----ILAPNNYPSKNLAPIALNAQKASSTPITMAYNQFL 416
DB	2228 LQNNMNNLENLNTIGSVQFVTCDFY--ANNIKSIYLD-----PILQNNNYFD 2277
QY	417 ELEKTKQLRLDT-----DQVYGNIAIYNFENGVRVDT-----GSNNSVLP 458
DB	2278 NMQLLNHNLNLNLSINVIDQNYLSYNNLNCINGNINKDKDVIIVIGPNNNTNQOI--P 2336
QY	459 QIOETTARIIFNGKOLNVLERRIAAVNPSDPLE-----TTKPDWTLKEALKIAFGP- 509

```
Db 2337 TI-ELDDTILKY--DUNLILNNVNVHVSUVEMLNTQNVQKLYNDIQENLHITNSLH 2393
QY 510 -NEPNGNIQYQOKD-----ITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKL 559
Db 2394 NNNNNNNMGLGCVLNNLSYLFNINSFNPNNKAYIYNNNNNNNNNNNNNNNNNNNN 2453
QY 560 NAK-----MNLIRDKRPHYDRNNIAGVADSUVKAEHREVNSTGLLANIDKIRKI 614
Db 2454 NVKNEIINNVSLEDTNKENEGD-----MDKNKYIINK--ENNKNNNNNIDII 2502
QY 615 LSGYIVEIEDTEGLKEVINDRYDMLNISLRQDGKTFIDFKYNDKLPFLYISNPNYKVV 674
Db 2503 SNGQNNNTNANTMEKNNLRDDDIYKQNMKRSNLSDFKKLDSE-----KKNKSNNI 2556
QY 675 ---YAVTKENT--IINPSENGDTSNGIKKILIFSKGYE 709
Db 2557 BEGARKMTNTTQLNLYSEN---RKKGLRDMSTYADKRVLE 2593

RESULT 15
Q81104
ID Q81104 PRELIMINARY; PRT; 3468 AA.
AC Q81104;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN Pf11_0371.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Hatt D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL; AE014841; AAN35955.1; -.
KW Hypothetical protein.
SQ SEQUENCE 3468 AA; 411913 MW; 506F7D62999BA7B1 CRC64;

Query Match 5.5%; Score 199.5; DB 5; Length 3468;
Best Local Similarity 19.0%; Pred. No. 0.81;
Matches 157; Conservative 130; Mismatches 263; Indels 275; Gaps 39;

QY 55 SENQYQSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEV---INKASNSNKIRLEKG 111
Db 160 SENY-----TKNDDKYNNSTNSGY-----NKEIEFLKNIKAKHSLKKEKN 202
QY 112 RLYQIK-----IQYQENPTEKGLDPK---LYW-----TDSQ-----KKEVISDNL 151
Db 203 INLLKXCDSDNNDLNLEFFKKNVHLNNDYRSEDLFHDNRVNNQNNKLDIIKDYIHDNI 262
QY 152 QLPKQKSNRKKKSTSGAGPTVPDRDNDGIPDSLEVEGYVDV-KNKRFTLSPWISNI 210
Db 263 MI-----ESNNDRNDKCNFKNTNPNYERHIIIVDTKNKGNNNNSNNFDS--ISNI 315
QY 211 HEKKGITKYSGPEKWTASDP-----YSDFE-KVTGRIDKNVSPPEARHPLVAAYPIHV- 264
Db 316 NKK-----ISYPINMYSNSEDKTLNQLNLDMSILSDSLKXNAYSFYSLS 360
QY 265 ---DMENIILSKNQDQSTQNTD--SETRTISKNTSRTHTS-----NSNS 305
```

```
Db 361 EKNSNNIGNKHKNMYRVDIVIDINISFNSINSLHDSMDNDNEFNSLCSFNSSS 420
QY 306 STVAIDHSLSLAGERTWAETMGL-----NTADTARLANANTRY-----VNTGTAPIYV 353
Db 421 KCVKDEMITQFVGNRKYVKTMDANDVYMKNSIEQHSINMESNHFKNQKNGSKEDAND 480
QY 354 LPTTSVLGKNQ-----TLATIKADE--NOLSOILAPNN----- 385
Db 481 NLNSSHILNNGKIGQVNSNLSDYLSIKKQSMNLSNNEALNINIVTNNNGSNNKX 540
QY 386 -----YYPKSNLAPIALNAQKQDASSTPITMYNQ-----FLELEKTKOL--RLDPTD 429
Db 541 NSNVKTSQVYVNPNDNIN--NMHQNLNLSYMKNSNLTNSGFKKIPKKNKNIISNIDPD 598
QY 430 Q-----VYGN-----ATYFENGSRVAVDGSNWSEVLPO-QEITAR----- 466
Db 599 NNIFKSYIKENVVKQOESNINHQVKEKNTYNDIEIKN-----NNIENNTQNTICNNFINT 554
QY 467 --IIFNGKDLNLVERRIAAVNPSPDPLETTK-----PDMTLKEA 502
Db 655 NDDIINKKKIKETIKYKIDSISLLNLSLNKLESNLSIMDRYTKKNYEEKFLDDVILDDS 714
QY 503 L-----KIAFGFNEPNCNLOYQOGKOITEFDNFQDQTSQNIKNQLAEL 545
Db 715 IFATSNELLQHSNYTTTNNHIFDNNNNNNNNINQKELQFONDYN---KETYNNNIMLSN 771
QY 546 NATNIYTVLDKIKLAKNNILIRDKFHYDRNNIAGVADSUVKAEHREVNSTEGLLL 605
Db 772 NAEMLF-----KISYSCNDLVLGKNEIILD--RNVENSKTE--- 806
QY 606 NIDKDIRKILSGYIVEIEDTEGLKE-----VINDRYDMLNLS--LRDQKTFIDF 654
Db 807 -----QYEEYQNKEDIILKYHKDDNIIDNNYNDVNIKKDCHLKMVDNQDNIKI 854
QY 655 KKYND---KLPL-----YISNPNYKVN-VYAVTKENTIINPSEN 689
Db 855 NKSQDSKKNKPINNNECNVHNKQFKINEIDHFKEENTLNVESKN 899

Search completed: May 3, 2004, 19:40:09
Job time : 36.4955 secs
```

```

QY 657 NDRYDMLNLSLRQDGKTFIDFKYNDKLPVYISNENYKVVAVTKENTIINPSENGDT 716
DB 641 NDRYDMLNLSLRQDGKTFIDFKYNDKLPVYISNENYKVVAVTKENTIINPSENGDT 700
QY 717 STNGIKKILIFSKGYEIG 735
DB 701 STNGIKKILIFSKGYEIG 719

RESULT 8
PCT-US94-01624-31
; Sequence 31, Application PC/TUS9401624
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; TITLE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
; STREET: Steuart Street Tower, 20th Floor, One Market
; STREET: Plaza
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01624
; FILING DATE: June 25, 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-115
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-01624-31

Query Match 94.8%; Score 3576; DB 5; Length 719;
Best Local Similarity 95.1%; Pred. No. 3.8e-249;
Matches 703; Conservative 4; Mismatches 8; Indels 24; Gaps 2

QY 1 EVKQENLLNSESSESSGCLGYFSDNLNFQAPMVVTSSTTGDLSIPSELENIENSENYQF 60
DB 1 EVKQENLLNSESSESSGCLGYFSDNLNFQAPMVVTSSTTGDLSIPSELENIENSENYQF 60
QY 61 QSAIWSGFIKVKKDEYTFATSDANHHVTMWDDQEVINKASNSKIRLEKGLRYQIKIY 120
DB 61 QSAIWSGFIKVKKDEYTFATSDANHHVTMWDDQEVINKASNSKIRLEKGLRYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVIGSDNLQLPELKQKSGNS----RKKSTASGPTVP 176
DB 121 QRENPTKGLDFKLYWTDSONKKEVIGSDNLQLPELKQKSGNSNTATINMQRGNFLOGPTVP 180
QY 177 DRDNDGIPDSLEVGSGYTDVKNKRTFLSPWLSNIHEKGLTKYKSSPEKWSSTASDPYSDF 236
DB 181 DRDNDGIPDSLEVGSGYTDVKNKRTFLSPWLSNIHEKGLTKYKSSPEKWSSTASDPYSDF 240

```

Db

421

LNAQDDFSPTITMNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV

480

Qy

481

LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL

540

Db

481

LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL

540

Qy

541

QYQKDIITEFDFNFQQTSONIKQLAELNATNIYTVLDKTKLNKAKNILLIRKRFHYDR

600

Db

541

QYQKDIITEFDFNFQQTSONIKQLAELNATNIYTVLDKTKLNKAKNILLIRKRFHYDR

600

Qy

601

NNIAGADESVVKEAHRVINSSTGLLNLDKDIRKILSGYIYEIEDTEGLKEVINDRY

660

Db

601

NNIAGADESVVKEAHRVINSSTGLLNLDKDIRKILSGYIYEIEDTEGLKEVINDRY

660

Qy

661

DMLNISSLRODQKTFIDFKKNDKPLIYISNPYKVNYYAVTKENTIIINPSENGDTSTNG

720

Db

661

DMLNISSLRODQKTFIDFKKNDKPLIYISNPYKVNYYAVTKENTIIINPSENGDTSTNG

720

Qy

721

IKKIL---IFSKKG 731

Db

721

IKKILKVLGKKG 734

RESULT 6

PCT-US94-01624-12

Sequence 12, Application PC/TUS9401624

GENERAL INFORMATION:

APPLICANT: Leppla, Stephen H.

APPLICANT: Klimpel, Kurt R.

APPLICANT: Arora, Naveen

APPLICANT: Singh, Yogendra

APPLICANT: Nichols, Peter J.

TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND

TITLE OF INVENTION: RELATED METHODS

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW

STREET: Stewart Street Tower, 20th Floor, One Market

STATE: Plaza

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/01624

FILING DATE: June 25, 1993

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 15280-115

TELEPHONE: (415) 543-9600

TELEFAX: (415) 543-5043

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 903 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-01624-12

Query Match 98.3%; Score 3708.5; DB 5; Length 903;

Best Local Similarity 98.8%; Pred. No. 1.5e-258;

Matches 725; Conservative 1; Mismatches 5; Indels 3; Gaps 1;

Db

1

EYKQENRLNSESQGLLYYFSDLNQAPMVVTSSTGTDLSPSELENIPSENQYF

60

Qy

61

QSAINSGFTKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY

120

Db

61

QSAINSGFTKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY

120

Qy

121

QRENPTKGLDFKLYWTDSONKEVITSSDNLQIPKQKSSNSRKRSTASGPTVPDRDN

180

Db

121

QRENPTKGLDFKLYWTDSONKEVITSSDNLQIPKQKSSNSRKRSTASGPTVPDRDN

180

Qy

181

DGIPDLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASDPSYDFEKT

240

Db

181

DGIPDLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASDPSYDFEKT

240

Qy

241

GRIDKVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT

300

Db

241

GRIDKVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT

300

Qy

301

SEVHGNAEVAHSPFDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL

360

Db

301

SEVHGNAEVAHSPFDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL

360

Qy

361

NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNYYPSKNLAPTA

420

Db

361

NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNYYPSKNLAPTA

420

Qy

421

LNAOKDASSTPTITMNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV

480

Db

421

LNAOKDASSTPTITMNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV

480

Qy

481

LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL

540

Db

481

LPOIQTETARIIFNGKDLNVLVERRIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPNGNL

540

Qy

541

QYQKDIITEFDFNFQQTSONIKQLAELNATNIYTVLDKTKLNKAKNILLIRKRFHYDR

600

Db

541

QYQKDIITEFDFNFQQTSONIKQLAELNATNIYTVLDKTKLNKAKNILLIRKRFHYDR

600

Qy

601

NNIAGADESVVKEAHRVINSSTGLLNLDKDIRKILSGYIYEIEDTEGLKEVINDRY

660

Db

601

NNIAGADESVVKEAHRVINSSTGLLNLDKDIRKILSGYIYEIEDTEGLKEVINDRY

660

Qy

661

DMLNISSLRODQKTFIDFKKNDKPLIYISNPYKVNYYAVTKENTIIINPSENGDTSTNG

720

Db

661

DMLNISSLRODQKTFIDFKKNDKPLIYISNPYKVNYYAVTKENTIIINPSENGDTSTNG

720

Qy

721

IKKIL---IFSKKG 731

Db

721

IKKILKVLGKKG 734

RESULT 7

US-08-082-849B-31

Sequence 31, Application US/08082849B

Patent No. 567274

GENERAL INFORMATION:

APPLICANT: Leppla, Stephen H.

APPLICANT: Klimpel, Kurt R.

APPLICANT: Arora, Naveen

APPLICANT: Singh, Yogendra

APPLICANT: Nichols, Peter J.

TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and

TITLE OF INVENTION: Related Methods

NUMBER OF SEQUENCES: 35

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk



INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-021-601-12

Query Match 98.3%; Score 3708.5; DB 1; Length 903;  
Best Local Similarity 98.8%; Pred. No. 1.5e-258;  
Matches 725; Conservative 1; Mismatches 5; Indels 3; Gaps 1;  
QY 1 EVKQENRLNESSESSQGLLYGYSFSLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNESSESSQGLLYGYSFSLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKYKSDRYTATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKYKSDRYTATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKRSTASGTPVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKRSTASGTPVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
Db 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFNGPENG 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFNGPENG 540  
QY 541 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYTVLQKILNKNQWILIRDKRFHYDR 600  
Db 541 QYQKDIETFDNFDDQTSQNIKNQLAELNATNIYTVLQKILNKNQWILIRDKRFHYDR 600  
QY 601 NNTAVGADSVVKEAREVINSSTEGLLNIDKDIRKILSGYVEIETEGLEKEVINDRY 660  
Db 601 NNTAVGADSVVKEAREVINSSTEGLLNIDKDIRKILSGYVEIETEGLEKEVINDRY 660  
QY 661 DMLNLSLRDQGTFFIDFKYNDKPLYSISNPNKYVAVTKENTIIINPSNGDTSTNG 720  
Db 661 DMLNLSLRDQGTFFIDFKYNDKPLYSISNPNKYVAVTKENTIIINPSNGDTSTNG 720  
QY 721 IKKIL----IFSKKG 731  
Db 721 IKKILKKVGLGKG 734

RESULT 5  
US-08-082-849B-12  
Sequence 12, Application US/08082849B  
Patent No. 5677274  
GENERAL INFORMATION:  
APPLICANT: Leppla, Stephen H.  
APPLICANT: Klmpel, Kurt R.  
APPLICANT: Arora, Naveen

APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
RELATED METHODS  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patencin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-12

Query Match 98.3%; Score 3708.5; DB 1; Length 903;  
Best Local Similarity 98.8%; Pred. No. 1.5e-258;  
Matches 725; Conservative 1; Mismatches 5; Indels 3; Gaps 1;  
QY 1 EVKQENRLNESSESSQGLLYGYSFSLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNESSESSQGLLYGYSFSLNFAQPMVVTSSITGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKYKSDRYTATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKYKSDRYTATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKRSTASGTPVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKRSTASGTPVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSPARPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
Db 421 LNAQKDASSPTIMYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480

QY 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720  
DB 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
DB 721 IKKILIFSKGYEIG 735

RESULT 3  
PCT-US94-01624-4  
; Sequence 4, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW  
; STREET: Stewart Street Tower, 20th Floor, One Market  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01624  
; FILING DATE: June 25, 1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-01624-4

Query Match 99.5%; Score 3753; DB 5; Length 735;  
Best Local Similarity 99.6%; Pred No. 7.2e-262;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60  
DB 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60  
QY 61 QSAINSGFTKVKSDYEYFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAINSGFTKVKSDYEYFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVYVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEY 240  
DB 181 DGIPDSLEVEGYTVYVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEY 240

QY 241 GRIDKNVSPPEARPLVAAYPIVHVDMENILSKNEDQSTONTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARPLVAAYPIVHVDMENILSKNEDQSTONTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NAMIRYNTGTAPIYNNVLTPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NAMIRYNTGTAPIYNNVLTPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTITMNNYNOLELEKTKQLRLDQVYGNATYNFENGVRVDTGSKWSEV 480  
DB 421 LNAQDDFSSPTITMNNYNOLELEKTKQLRLDQVYGNATYNFENGVRVDTGSKWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540  
DB 481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENG 540  
QY 541 QYQKDIETDFNFDQOTSONIKQNLAEALNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600  
DB 541 QYQKDIETDFNFDQOTSONIKQNLAEALNATNIYTVLDKIKLNKAKNLIIRDKRPHYDR 600  
QY 601 NNIAVGADSVKAEHREVINSSTGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
DB 601 NNIAVGADSVKAEHREVINSSTGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720  
DB 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNYKVNVAVTKENTIIINPSENGDSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
DB 721 IKKILIFSKGYEIG 735

RESULT 4  
US-08-021-601-12  
; Sequence 12, Application US/08021601  
; Patent No. 5591631  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Nichols, Peter J.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; RELATED METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Needle & Rosenberg, P.C.  
; STREET: 133 Carnegie Way, Suite 400  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,601  
; FILING DATE: 19930212  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 1414.057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880

Db 1 EVKQENRLNSESQGLLYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGNSNRKRSTSGPTVPDRDN 180  
Db 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGNSNRKRSTSGPTVPDRDN 180  
QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPSDFEKT 240  
Db 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPSDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKQDASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVFNENGRVVDTSNWSSEV 480  
Db 421 LNAQKQDASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVFNENGRVVDTSNWSSEV 480  
QY 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKBAKIAFGFNEPENG 540  
Db 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKBAKIAFGFNEPENG 540  
QY 541 QYQKQDITEFDNFDOOTSQNKQLAELNATNIVTLDKIKLAKXNMLIRDKRFHYDR 600  
Db 541 QYQKQDITEFDNFDOOTSQNKQLAELNATNIVTLDKIKLAKXNMLIRDKRFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660  
QY 661 DMLNSSLQDQKTFIDFKYNDKPLYSNPNYKVNVTAKNTIINPSENGDTSTNG 720  
Db 661 DMLNSSLQDQKTFIDFKYNDKPLYSNPNYKVNVTAKNTIINPSENGDTSTNG 720  
QY 721 IKKILFSKKGVEIG 735  
Db 721 IKKILFSKKGVEIG 735

## RESULT 2

US-08-082-849B-4  
; Sequence 4, Application US/08082849B  
; Patent No. 5677274  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Kimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
; TITLE OF INVENTION: Related Methods  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,849B  
; FILING DATE: 25-JUN-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/021,601  
; FILING DATE: 12-FEB-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-161-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-082-849B-4

Query Match 99.5%; Score 3753; DB 1; Length 735;  
Best Local Similarity 99.6%; Pred. No. 7.2e-262;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESQGLLYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLLYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGNSNRKRSTSGPTVPDRDN 180  
Db 121 QRENTEKGLDPLKLYWDSQNKKEVSSDNLOLPELKOKGNSNRKRSTSGPTVPDRDN 180  
QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPSDFEKT 240  
Db 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPSDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKQDASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVFNENGRVVDTSNWSSEV 480  
Db 421 LNAQKQDASSTPTIMYVNFQLEKTKQLRLDQVYGNATYVFNENGRVVDTSNWSSEV 480  
QY 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKBAKIAFGFNEPENG 540  
Db 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKBAKIAFGFNEPENG 540  
QY 541 QYQKQDITEFDNFDOOTSQNKQLAELNATNIVTLDKIKLAKXNMLIRDKRFHYDR 600  
Db 541 QYQKQDITEFDNFDOOTSQNKQLAELNATNIVTLDKIKLAKXNMLIRDKRFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 15.0986 Seconds  
(without alignments)  
2513.152 Million cell updates/sec

Title: US-09-848-909A-13

Perfect score: 3772

Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILIFSKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/2/aaa/5A\_COMB.pep:\*

2: /cgn2\_6/prodata/2/aaa/5B\_COMB.pep:\*

3: /cgn2\_6/prodata/2/aaa/6A\_COMB.pep:\*

4: /cgn2\_6/prodata/2/aaa/6B\_COMB.pep:\*

5: /cgn2\_6/prodata/2/aaa/PCFUS\_COMB.pep:\*

6: /cgn2\_6/prodata/2/aaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3753	99.5	735	1	US-08-021-601-4
2	3753	99.5	735	1	US-08-082-849B-4
3	3753	99.5	735	5	PCF-US94-01624-4
4	3708.5	98.3	903	1	US-08-021-601-12
5	3708.5	98.3	903	1	US-08-082-849B-12
6	3708.5	98.3	903	5	PCF-US94-01624-12
7	3576	94.8	719	1	US-08-082-849B-31
8	3576	94.8	719	5	PCF-US94-01624-31
9	1275	33.8	288	4	US-09-273-839A-8
10	767	20.3	881	3	US-08-960-780-32
11	767	20.3	881	3	US-09-073-898-32
12	767	20.3	881	4	US-09-307-106-8
13	767	20.3	881	4	US-09-850-351A-32
14	765	20.3	884	1	US-08-471-033-5
15	765	20.3	884	2	US-08-471-044-5
16	765	20.3	884	2	US-08-463-483A-5
17	765	20.3	884	2	US-08-471-046A-5
18	765	20.3	884	2	US-08-470-566B-5
19	765	20.3	884	2	US-08-469-334-5
20	765	20.3	884	2	US-09-300-529-5
21	765	20.3	1346	1	US-08-471-033-23
22	765	20.3	1346	2	US-08-471-044-23
23	765	20.3	1346	2	US-08-463-483A-23
24	765	20.3	1346	2	US-08-471-046A-23
25	765	20.3	1346	2	US-08-470-566B-23
26	765	20.3	1346	2	US-08-469-334-23
27	765	20.3	1346	3	US-09-300-529-23

28	764	20.3	852	1	US-08-471-033-36	Sequence 36, Appl
29	764	20.3	852	2	US-08-471-044-36	Sequence 36, Appl
30	764	20.3	852	2	US-08-463-483A-36	Sequence 36, Appl
31	764	20.3	852	2	US-08-471-046A-36	Sequence 36, Appl
32	764	20.3	852	2	US-08-470-566B-36	Sequence 36, Appl
33	764	20.3	852	2	US-08-469-334-36	Sequence 36, Appl
34	764	20.3	852	3	US-09-300-529-36	Sequence 36, Appl
35	764	20.3	1338	1	US-08-471-033-50	Sequence 50, Appl
36	764	20.3	1338	2	US-08-471-044-50	Sequence 50, Appl
37	764	20.3	1338	2	US-08-463-483A-50	Sequence 50, Appl
38	764	20.3	1338	2	US-08-471-046A-50	Sequence 50, Appl
39	764	20.3	1338	2	US-08-470-566B-50	Sequence 50, Appl
40	764	20.3	1338	2	US-08-469-334-50	Sequence 50, Appl
41	764	20.3	1338	3	US-09-300-529-50	Sequence 50, Appl
42	754.5	20.0	784	3	US-09-371-913A-7	Sequence 7, Appl
43	754.5	20.0	784	4	US-09-967-805-7	Sequence 7, Appl
44	745.5	19.8	860	4	US-08-307-106-48	Sequence 48, Appl
45	735.5	19.5	834	1	US-08-471-033-21	Sequence 21, Appl

#### ALIGNMENTS

RESULT 1  
US-08-021-601-4  
; Sequence 4, Application US/08021601  
; Patent No. 5591631  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Nichols, Peter J.  
; APPLICANT: Arora, Naveen  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Needle & Rosenberg, P.C.  
; STREET: 133 Carnegie Way, Suite 400  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,601  
; FILING DATE: 19930212  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 1414.057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-021-601-4

Query Match 99.5%; Score 3753; DB 1; Length 735;  
Best Local Similarity 99.6%; Pred. No. 7.2e-262;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESQGLLYGFFDLNFOAPWVTSSTTGGDLSPSSLENTIPSENOYF 60  
|||||

QY 661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 720  
DB 661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 15  
AAV56959 standard; protein; 736 AA.  
XX AAY56959;  
AC  
XX 25-APR-2000 (first entry)  
XX B. anthracis MAT-PA protein.  
DE  
XX Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;  
KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.  
XX  
OS Bacillus anthracis.  
XX  
XX WO200002522-A2.  
XX 20-JAN-2000.  
XX 09-JUL-1999; 99WO-US015568.  
XX 10-JUL-1998; 98US-0092416P.  
XX (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.  
XX Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;  
PI WPI; 2000-182165/16.  
XX N-PSDB; AAZ56875.  
PT Recombinant DNA construct useful as vaccines for anthrax, in producing  
PT host cells for analyzing the drugs and agents inhibiting anthrax.  
XX  
PS Disclosure; Page 34; 35pp; English.  
XX  
XX The invention provides a recombinant DNA construct that comprises a  
XX vector and at least one nucleic acid (or its fragment) encoding a  
XX combination of Bacillus anthracis proteins, selected from protective  
XX antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA  
XX with its secretory signals replaced with those of tissue plasminogen  
XX activator) and PA63 (63 KDa PA). The construct is useful as a DNA vaccine  
XX for anthrax and in producing infectious alpha virus particles. These  
XX particles, expressing the B. anthracis proteins are useful also as  
XX vaccines for anthrax. Host cells transformed with the construct are  
XX useful for analyzing the effectiveness of drugs and agents that inhibit  
XX anthrax or B. anthracis proteins. The present sequence represents a B.  
XX anthracis MAT-PA protein  
SQ Sequence 736 AA;

Query Match 99.5%; Score 3753; DB 3; Length 736;  
Best Local Similarity 99.6%; Pred. No. 1.6e-247;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYPFDLNFQPMVVTSTTGDLSIPSELENIPSENQYF 60  
DB 2 EVKQENRLNSESQGLLYPFDLNFQPMVVTSTTGDLSIPSELENIPSENQYF 61  
QY 61 QSATWSGPIKVKSDYEYTFATSDNHVTWDDQEVINKASNSNKIRLEKGRLYQIKIY 120  
DB 62 QSATWSGPIKVKSDYEYTFATSDNHVTWDDQEVINKASNSNKIRLEKGRLYQIKIY 121  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNILQLPKQSSNSRKRKTSAGPTVPDRN 180

DB 122 QRENPTKGLDFKLYWTDSONKKEVISSDNILQLPKQSSNSRKRKTSAGPTVPDRN 181  
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPSDFEYV 240  
DB 182 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPSDFEYV 241  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDNENILSKNVEDQSTQNTDSETRTISKNITSRTHT 300  
DB 242 GRIDKNVSPARHPLVAAYPIVHVDNENILSKNVEDQSTQNTDSETRTISKNITSRTHT 301  
QY 301 SEVHGNAEVHASFDFIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 302 SEVHGNAEVHASFDFIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 361  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSOILAPNNYPPSKNLAPIA 420  
DB 362 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSOILAPNNYPPSKNLAPIA 421  
QY 421 LNAQKDASSPTITWYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVVDTSNWSSEV 480  
DB 422 LNAQKDASSPTITWYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVVDTSNWSSEV 481  
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAGFNPNPGL 540  
DB 482 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIAGFNPNPGL 541  
QY 541 QYQCKDITEFDNFDQOTSONIKNQLAELNATNIYTVLDKIKLNAKXNILLIRKRPHYDR 600  
DB 542 QYQCKDITEFDNFDQOTSONIKNQLAELNATNIYTVLDKIKLNAKXNILLIRKRPHYDR 601  
QY 601 NNIAVGADES VVKEAHEVINSTEGLLLNIDKIDIRKILSGYIVEIEDTEGLKEVINDRY 660  
DB 602 NNIAVGADES VVKEAHEVINSTEGLLLNIDKIDIRKILSGYIVEIEDTEGLKEVINDRY 661  
QY 661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 720  
DB 662 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKEHTIINPSENGDTSTNG 721  
QY 721 IKKILIFSKKGYEIG 735  
DB 722 IKKILIFSKKGYEIG 736

Search completed: May 3, 2004, 19:36:03  
Job time : 52.1208 secs

Db 61 QSAIWSGFIKVKSDYEYFATSADNHVTWVDDQEVINKASNSKIRLEKGLQIKIY 120  
QY 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOPELKOKSNSRKRSTSGPTVPDRN 180  
Db 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOPELKOKSNSRKRSTSGPTVPDRN 180  
QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYPYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYPYPSKNLAPIA 420  
QY 421 LNAOKDASSTPIITWVYNOFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480  
Db 421 LNAOKDASSTPIITWVYNOFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540  
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNITVLDKIKLNKXNILLIRDKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNITVLDKIKLNKXNILLIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIIVEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIIVEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKNDKPLIYSNPNYKVNVAVTXENTINPSENGDTSTNG 720  
Db 661 DMLNSSLRQDGKTFIDFKKNDKPLIYSNPNYKVNVAVTXENTINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYBIG 735  
Db 721 IKKILIFSKKGYBIG 735

XX PR 04-MAY-2000; 2000US-0201800P.  
XX PA (HARD ) HARVARD COLLEGE.  
XX PI Collier RJ, Sellman BR;  
XX DR WPI; 2002-017725/02.  
XX PT Protecting humans against anthrax using mutant B groups (anthrax  
XX PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX PT anthracis.  
XX PS Claim 4; Page: 77pp; English.  
XX CC The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAMS1483)  
XX SQ Sequence 735 AA;  
Query Match 99.5%; Score 3753; DB 5; Length 735;  
Best Local Similarity 99.6%; Pred. No. 1.6e-247;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENLLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
Db 1 EVKQENLLNESSSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDYEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRUYQIKIY 120  
Db 61 QSAIWSGFIKVKSDYEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRUYQIKIY 120  
QY 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOPELKOKSNSRKRSTSGPTVPDRN 180  
Db 121 QRENTEKGLDFKLWYTSQNKKEVSSDNLOPELKOKSNSRKRSTSGPTVPDRN 180  
QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYPYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNYPYPSKNLAPIA 420  
QY 421 LNAOKDASSTPIITWVYNOFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480  
Db 421 LNAOKDASSTPIITWVYNOFLEKTKQLRLDQVYGNATYVNFENGVRVDTGSNWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540  
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNITVLDKIKLNKXNILLIRDKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNITVLDKIKLNKXNILLIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIIVEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIIVEIDTEGLKEVINDRY 660

RESULT 14  
AAMS1499  
ID AAMS1499 standard; protein; 735 AA.  
XX AC AAMS1499;  
XX DE Anthrax PA F427D mutant.  
XX DE 01-FEB-2002 (first entry)  
XX DE Anthrax PA F427D mutant.  
XX KW Anthrax; PA: protective antigen; antibacterial; pore-forming toxin;  
XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.  
XX OS Bacillus anthracis.  
XX OS Synthetic.  
XX FT Key Location/Qualifiers  
XX FT Misc-difference 427 /note= "Wild-type Phe substituted by Asp"  
XX FT WO200182788-A2.  
XX PD 08-NOV-2001.  
XX PP 04-MAY-2001; 2001WO-US014372.

```
PA (HARD ) HARVARD COLLEGE.
XX
PI Collier RJ, Sellman BR;
XX
XX WPI; 2002-017725/02.
DR
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX
XX Claim 4; Page; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
XX Sequence 735 AA;
XX
XX Query Match 99.5%; Score 3753; DB 5; Length 735;
XX Best Local Similarity 99.6%; Pred. No. 1.6e-247;
XX Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 EVKQENRLNESSSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFTKVKSDYEYFATSADNHVTMWDDQEVINKASNSKIRLEKRLYQIKQY 120
Db 61 QSAIWSGFTKVKSDYEYFATSADNHVTMWDDQEVINKASNSKIRLEKRLYQIKQY 120
QY 121 QRENPEKGLDFKLYWTDSONKKEVISDNLOLPELKQKSNRSKRKSTASGTPVDRDN 180
Db 121 QRENPEKGLDFKLYWTDSONKKEVISDNLOLPELKQKSNRSKRKSTASGTPVDRDN 180
QY 181 DGIPTDSEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGIPTDSEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
QY 301 SEVHGNAEVHAFDIDGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAFDIDGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNYYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNYYPSKNLAPIA 420
QY 421 LNAQKDASTPTIMYNNQFLEKTKQLDLDDQVYGNATYNNPENGVRVDTGNSWSEV 480
Db 421 LNAQKDASTPTIMYNNQFLEKTKQLDLDDQVYGNATYNNPENGVRVDTGNSWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATGFPNPNGL 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATGFPNPNGL 540
QY 541 QYQKDIITFDNFDOOTSQNKQLAELNATNIYTVLDKIKLNAMILLRDKGFHYDR 600
Db 541 QYQKDIITFDNFDOOTSQNKQLAELNATNIYTVLDKIKLNAMILLRDKGFHYDR 600
QY 601 NNTAVGADESVKAEHREVINSGTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNTAVGADESVKAEHREVINSGTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNISLSDQDGTFTDFKKNYNDKPLXISNPNYKVVYAVTKENTINPSENGTSTNG 720
Db 661 DMLNISLSDQDGTFTDFKKNYNDKPLXISNPNYKVVYAVTKENTINPSENGTSTNG 720
Db 661 DMLNISLSDQDGTFTDFKKNYNDKPLXISNPNYKVVYAVTKENTINPSENGTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
XX
XX RESULT 13
XX AAM51494
XX ID AAM51494 standard; protein; 735 AA.
XX
XX AC AAM51494;
XX
XX DT 01-FEB-2002 (first entry)
XX
XX DE Anthrax PA mutant K395D/K397D/D425K/D426K.
XX
XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutoin.
XX
XX KW Bacillus anthracis.
XX
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 395 /note= "Wild-type Lys substituted by Asp"
XX Misc-difference 397 /note= "Wild-type Lys substituted by Asp"
XX Misc-difference 425 /note= "Wild-type Asp substituted by Lys"
XX Misc-difference 426 /note= "Wild-type Asp substituted by Lys"
XX
XX WC200182788-A2.
XX
XX PD 08-NOV-2001.
XX
XX PF 04-MAY-2001; 2001WO-US014372.
XX
XX PR 04-MAY-2000; 2000US-0201800P.
XX
XX PA (HARD ) HARVARD COLLEGE.
XX
XX PI Collier RJ, Sellman BR;
XX
XX DR WPI; 2002-017725/02.
XX
XX PS Claim 4; Page; 77pp; English.
XX
XX CC The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
XX Sequence 735 AA;
XX
XX Query Match 99.5%; Score 3753; DB 5; Length 735;
XX Best Local Similarity 99.6%; Pred. No. 1.6e-247;
XX Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 EVKQENRLNESSSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFTKVKSDYEYFATSADNHVTMWDDQEVINKASNSKIRLEKRLYQIKQY 120
Db 61 QSAIWSGFTKVKSDYEYFATSADNHVTMWDDQEVINKASNSKIRLEKRLYQIKQY 120
QY 121 QRENPEKGLDFKLYWTDSONKKEVISDNLOLPELKQKSNRSKRKSTASGTPVDRDN 180
Db 121 QRENPEKGLDFKLYWTDSONKKEVISDNLOLPELKQKSNRSKRKSTASGTPVDRDN 180
QY 181 DGIPTDSEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGIPTDSEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKWKSTASDPYDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Db 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
QY 301 SEVHGNAEVHAFDIDGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAFDIDGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNYYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILLAPNYYPSKNLAPIA 420
QY 421 LNAQKDASTPTIMYNNQFLEKTKQLDLDDQVYGNATYNNPENGVRVDTGNSWSEV 480
Db 421 LNAQKDASTPTIMYNNQFLEKTKQLDLDDQVYGNATYNNPENGVRVDTGNSWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATGFPNPNGL 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATGFPNPNGL 540
QY 541 QYQKDIITFDNFDOOTSQNKQLAELNATNIYTVLDKIKLNAMILLRDKGFHYDR 600
Db 541 QYQKDIITFDNFDOOTSQNKQLAELNATNIYTVLDKIKLNAMILLRDKGFHYDR 600
QY 601 NNTAVGADESVKAEHREVINSGTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
Db 601 NNTAVGADESVKAEHREVINSGTEGLLNIDKDKIRKILSGYIVIEDTEGLKEVINDRY 660
QY 661 DMLNISLSDQDGTFTDFKKNYNDKPLXISNPNYKVVYAVTKENTINPSENGTSTNG 720
Db 661 DMLNISLSDQDGTFTDFKKNYNDKPLXISNPNYKVVYAVTKENTINPSENGTSTNG 720
```

Db 541 QYQKDIETDFDNFQDOTSONIKNQLAELNATNIYTVLDKIKLNKMWILIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735

RESULT 11  
AAM51483  
ID AAM51483 standard; protein; 735 AA.  
XX AC AAM51483;  
DT 01-FEB-2002 (first entry)  
DE Anthrax PA protein.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.  
XX  
OS Bacillus anthracis.  
XX WO200182788-A2.  
XX  
PD 08-NOV-2001.  
XX  
PF 04-MAY-2001; 2001WO-US014372.  
XX  
PR 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
DR N-PSDB; AA199904.  
XX  
XX Protecting humans against anthrax using mutant B groups (anthrax  
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
PT anthracis.  
XX  
PS Disclosure; Fig 13; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. The present sequence is that of the  
CC anthrax PA protein  
XX  
SQ Sequence 735 AA;

Query Match 99.5%; Score 3753; DB 5; Length 735;  
Best Local Similarity 99.6%; Pred. No. 1.6e-247;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVVTSSTGDLSTPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVVTSSTGDLSTPSSSELENIPSENOYF 60  
QY 61 QSAINSGFIKVKSDBYTTFATSDADNHTVWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAINSGFIKVKSDBYTTFATSDADNHTVWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVTSDDNLOQLPELKQSSNSRKKRSTSGAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSDDNLOQLPELKQSSNSRKKRSTSGAGTVPDRDN 180  
QY 181 DGI:PDLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDPEKVT 240  
Db 181 DGI:PDLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDPEKVT 240  
QY 241 GRIDKNVSPPEARPIVAAPYIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTGRHT 300  
Db 241 GRIDKNVSPPEARPIVAAPYIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSTGRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAFSSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAFSSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSQILAPNNYPPSKNLAPIA 420  
QY 421 LNAQKDCASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAATYFNENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKDCASSTPITMNYNQFLELEKTKQLRLDTPQVYGNIAATYFNENGRVVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENGNL 540  
QY 541 QYQKDIETDFDNFQDOTSONIKNQLAELNATNIYTVLDKIKLNKMWILIRDKRFHYDR 600  
Db 541 QYQKDIETDFDNFQDOTSONIKNQLAELNATNIYTVLDKIKLNKMWILIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHREVINSTEGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLIYISNPNYKVNVAVTKENTIINPSENGDTSNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735

RESULT 12  
AAM51488  
ID AAM51488 standard; protein; 735 AA.  
XX AC AAM51488;  
DT 01-FEB-2002 (first entry)  
DE Anthrax PA mutant D425A.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX Bacillus anthracis.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 425 /note="Wild-type Asp substituted by Ala"  
FT  
XX WO200182788-A2.  
XX 08-NOV-2001.  
XX 04-MAY-2001; 2001WO-US014372.  
XX 04-MAY-2000; 2000US-0201800P.  
XX



QY	1	EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQYF	60
Db	1	EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQYF	60
QY	61	QSAIWSGFIKVKKSDGYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY	120
Db	61	QSAIWSGFIKVKKSDGYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY	120
QY	121	QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNSRKRSTAGPTVPDRDN	180
Db	121	QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNSRKRSTAGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWTASDPYDFEKT	240
Db	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWTASDPYDFEKT	240
QY	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
Db	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
QY	301	SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Db	361	NANIRYVNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKDASSTPITMNYNQFLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGNSNWEV	480
Db	421	LNAQKDASSTPITMNYNQFLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGNSNWEV	480
QY	481	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG	540
Db	481	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG	540
QY	541	QYQKDIETFDNFDOQTSONIKNQLAELNATNITVLDKIKNAKNNILIRDKRPHYDR	600
Db	541	QYQKDIETFDNFDOQTSONIKNQLAELNATNITVLDKIKNAKNNILIRDKRPHYDR	600
QY	601	NNIAVGADESUVKEAHEVINSSTEGLLINIDKDIRKILSGYVIEIDTEGLKEVINDRY	660
Db	601	NNIAVGADESUVKEAHEVINSSTEGLLINIDKDIRKILSGYVIEIDTEGLKEVINDRY	660
QY	661	DMLNISLRQDGKTFIDFKKYNDKPLXYISNPNYKVNVAVTKNTIINPSENGDTSTNG	720
Db	661	DMLNISLRQDGKTFIDFKKYNDKPLXYISNPNYKVNVAVTKNTIINPSENGDTSTNG	720
QY	721	IKKILIFSKKGYEIG 735	
Db	721	IKKILIFSKKGYEIG 735	
RESULT 10			
AAR60179			
XX	ID	AAR60179 standard; protein; 735 AA.	
XX	AC	AAR60179;	
XX	CC		
XX	DT	25-MAR-2003 (revised)	
XX	DT	03-APR-1995 (first entry)	
XX	DE	Protective antigen of Bacillus anthracis.	
XX	KW	Anthrax; Bacillus anthracis; fusion protein; protective antigen;	
XX	KW	protective antigen; cell killing; targeting; toxin;	
XX	KW	intracellular; HIV; human immunodeficiency virus; toxin.	
XX	OS	Bacillus anthracis.	
XX	PN	W09418332-A2.	
XX	XX		

PD	19-AUG-1994.		
XX	14-FEB-1994;	94WO-US001624.	
XX	12-FEB-1993;	93US-00021601.	
PR	25-JUN-1993;	93US-00082849.	
XX	(USSH )	US DEPT HEALTH & HUMAN SERVICES.	
XX	Leppla SH,	Klimpel K, Arora N, Singh Y, Nichols PJ;	
PI	WPI;	1994-279753/34.	
DR	N-PSDB;	AAQ70180.	
XX			
XX			
PT			
PT			
XX			
PS			
XX			
CC			
CC			
CC			
CC			
CC			
CC			
CC			
XX			
SQ			

Query Match	99.5%;	Score 3753;	DB 2;	Length 735;
Best Local Similarity	99.6%;	Pred. No. 1.6e-247;		
Matches	732;	Conservative 0;	Mismatches 3;	Indels 0; Gaps 0;

QY	1	EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQYF	60
Db	1	EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQYF	60
QY	61	QSAIWSGFIKVKKSDGYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY	120
Db	61	QSAIWSGFIKVKKSDGYTFATSDNHNVTWVDDQEVINKASNSKIRLEKGLYQIKIY	120
QY	121	QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNSRKRSTAGPTVPDRDN	180
Db	121	QRENTEKGLDFKLYWTDSONKKEVISSDNILQLPELKQKSSNSRKRSTAGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWTASDPYDFEKT	240
Db	181	DGIPDSLEVEGYTVVKNKRTFLSPWISNIHKKGLTKYKSSPEKWTASDPYDFEKT	240
QY	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
Db	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
QY	301	SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Db	361	NANIRYVNTGTAPIYNNVLTPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKDASSTPITMNYNQFLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGNSNWEV	480
Db	421	LNAQKDASSTPITMNYNQFLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGNSNWEV	480
QY	481	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG	540
Db	481	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG	540
QY	541	QYQKDIETFDNFDOQTSONIKNQLAELNATNITVLDKIKNAKNNILIRDKRPHYDR	600

FF	Key	Location/Qualifiers	
PT	Misc-difference 427	/note= "Wild-type Phe substituted by Lys"	
XX			
PN	WO200182788-A2.		
XX			
PD	08-NOV-2001.		
XX			
PF	04-MAY-2001; 2001WO-US014372.		
XX			
PR	04-MAY-2000; 2000US-0201800P.		
XX			
PA	(HARD ) HARVARD COLLEGE.		
XX			
PI	Collier RJ, Sellman BR;		
XX			
DR	WPI; 2002-017725/02.		
XX			
PT	Protecting humans against anthrax using mutant B groups (anthrax		
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus		
XX	anthracis.		
PS	Claim 4; Page; 77pp; English.		
XX			
CC	The invention relates to antibacterial agents comprising mutant forms of		
CC	pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants		
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B		
CC	moiety is anthrax protective antigen (PA) and using these mutants or		
CC	compositions of them for protecting against Bacillus anthracis infections		
CC	in humans, especially as vaccines. Note: The present sequence is not		
CC	given in the specification but is derived from the Bacillus anthracis		
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)		
XX			
SQ	Sequence 735 AA;		
	Query Match 99.5%; Score 3754; DB 5; Length 735;		
	Best Local Similarity 99.6%; Pred. No. 1.4e-247;		
	Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		
QY	1	EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTTGDLSPSSSELENIPSENQYF 60	
DB	1	EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTTGDLSPSSSELENIPSENQYF 60	
QY	61	QSAINSGFTKVKSDYEYFATGADNHVTWVDQEVINKASNSKIRLEKGLVQIKQY 120	
DB	61	QSAINSGFTKVKSDYEYFATGADNHVTWVDQEVINKASNSKIRLEKGLVQIKQY 120	
QY	121	QRENPEKGLDFKLYWTDSONKKEVSSDNLOLPELKOKSSNSRKRSTSGPTVPDRDN 180	
DB	121	QRENPEKGLDFKLYWTDSONKKEVSSDNLOLPELKOKSSNSRKRSTSGPTVPDRDN 180	
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWSASDPYSDFEKT 240	
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGITKYKSSPEKWSASDPYSDFEKT 240	
QY	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300	
DB	241	GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300	
QY	301	SEVHGNAEYHAFDFDGGSVSAGFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
DB	301	SEVHGNAEYHAFDFDGGSVSAGFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360	
QY	361	NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420	
DB	361	NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420	
QY	421	LNAQKQASPTPTMNTNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGNSWSEV 480	
DB	421	LNAQKQASPTPTMNTNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGNSWSEV 480	
QY	481	LPOIQETTARIIFNGKDLNVERRIAANVPSPDLETTKPDMTLKEALKIAPGFNEPNGNL 540	

DB	481	LPOIQETTARIIFNGKDLNVERRIAANVPSPDLETTKPDMTLKEALKIAPGFNEPNGNL 540	
QY	541	OYQOKDITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAKQNLILIRDKRPHYDR 600	
DB	541	OYQOKDITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAKQNLILIRDKRPHYDR 600	
QY	601	NNIAGVADSVVKEAHRVINSSTEGLLNIDKIDRKILSGYIVIEIDTEGLKEVINDRY 660	
DB	601	NNIAGVADSVVKEAHRVINSSTEGLLNIDKIDRKILSGYIVIEIDTEGLKEVINDRY 660	
QY	661	DLNLISSLRQDGKTFIDFKKYNDKLPYISNPNYKVNKYAVTKENTLIINPSENGDSTNG 720	
DB	661	DLNLISSLRQDGKTFIDFKKYNDKLPYISNPNYKVNKYAVTKENTLIINPSENGDSTNG 720	
QY	721	IKKILIFSKKGYEIG 735	
DB	721	IKKILIFSKKGYEIG 735	
	RESULT 9		
	AAM51489		
ID	AAM51489 standard; protein; 735 AA.		
XX			
AC	AAM51489;		
XX			
DT	01-FEB-2002 (first entry)		
XX			
DE	Anthrax PA mutant D425N.		
XX			
KW	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;		
KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.		
XX	Bacillus anthracis.		
OS	Synthetic.		
XX			
PH	Key Location/Qualifiers		
FT	Misc-difference 425		
FT	/note= "Wild-type Asp substituted by Asn"		
XX			
FN	WO200182788-A2.		
XX			
PD	08-NOV-2001.		
XX			
PF	04-MAY-2001; 2001WO-US014372.		
XX			
PR	04-MAY-2000; 2000US-0201800P.		
XX			
PA	(HARD ) HARVARD COLLEGE.		
XX			
PI	Collier RJ, Sellman BR;		
XX			
DR	WPI; 2002-017725/02.		
XX			
PT	Protecting humans against anthrax using mutant B groups (anthrax		
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus		
XX	anthracis.		
PS	Claim 4; Page; 77pp; English.		
XX			
CC	The invention relates to antibacterial agents comprising mutant forms of		
CC	pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants		
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B		
CC	moiety is anthrax protective antigen (PA) and using these mutants or		
CC	compositions of them for protecting against Bacillus anthracis infections		
CC	in humans, especially as vaccines. Note: The present sequence is not		
CC	given in the specification but is derived from the Bacillus anthracis		
CC	wild-type PA protein sequence shown in figure 13 (AAM51483)		
XX			
SQ	Sequence 735 AA;		
	Query Match 99.5%; Score 3754; DB 5; Length 735;		
	Best Local Similarity 99.6%; Pred. No. 1.4e-247;		
	Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;		

481 LPQIQETTARIIFNGKDLNVLVERRIAAVNPSPDPLETTKPDMTLKEALKIAFGFNEPENGNL 540  
481 LPQIQETTARIIFNGKDLNVLVERRIAAVNPSPDPLETTKPDMTLKEALKIAFGFNEPENGNL 540  
541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660  
601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660  
661 DMLNLSLRQDGTFTDFPKYNDKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
661 DMLNLSLRQDGTFTDFPKYNDKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
721 IKKILIFSCKGYEIG 735  
721 IKKILIFSCKGYEIG 735

RESULT 7  
AAM51487  
ID AAM51487 standard; protein; 735 AA.  
XX AC AAM51487;

01-FEB-2002 (first entry)  
Anthrax PA mutant K397Q.  
XX  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX  
XX Bacillus anthracis.  
XX Synthetic.

Key Location/Qualifiers  
FT Misc-difference 397 /note= "Wild-type Lys substituted by Gln"  
XX  
XX WO200182788-A2.  
XX  
XX 08-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US014372.  
XX  
XX 04-MAY-2000; 2000US-0201800P.  
XX  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Collier RJ, Sellman BR;  
XX WPT; 2002-017725/02.  
XX

Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.  
Claim 4; Page; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of  
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX moiety is anthrax protective antigen (PA) and using these mutants or  
XX compositions of them for protecting against Bacillus anthracis infections  
XX in humans, especially as vaccines. Note: The present sequence is not  
XX given in the specification but is derived from the Bacillus anthracis  
XX wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
XX Sequence 735 AA;

Query Match 99.5%; Score 3754; DB 5; Length 735;  
Best Local Similarity 99.6%; Pred. No. 1.4e-247; Indels 0; Gaps 0;  
Matches 732; Conservative 0; Mismatches 3;  
QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFAQPMVYTSSTTGDLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLLYYFSDLNFAQPMVYTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWGSFIVKVKSDYETFTATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
DB 61 QSAIWGSFIVKVKSDYETFTATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLWYTDSONKKEVISSDNLQIPELKQKSSNSRKRSTTSAGTTPVDRDN 180  
DB 121 QRENPTKGLDFKLWYTDSONKKEVISSDNLQIPELKQKSSNSRKRSTTSAGTTPVDRDN 180  
QY 181 DGIPDSLEVEGYTVDKNRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDYSDEKVT 240  
DB 181 DGIPDSLEVEGYTVDKNRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDYSDEKVT 240  
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300  
DB 241 GRIDKNVSPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300  
QY 301 SEVHGNAEVAHSPFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVAHSPFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLTPTSLVLCNKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYNNVLTPTSLVLCNKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTITMWNQFLEKTKQLRLDQVYGNATYVNFNGRVVRVDTGSWSEV 480  
DB 421 LNAQKDASSPTITMWNQFLEKTKQLRLDQVYGNATYVNFNGRVVRVDTGSWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVLVERRIAAVNPSPDPLETTKPDMTLKEALKIAFGFNEPENG 540  
DB 481 LPQIQETTARIIFNGKDLNVLVERRIAAVNPSPDPLETTKPDMTLKEALKIAFGFNEPENG 540  
QY 541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
DB 541 QYQKDIITEFDNFQOQTSQNTKQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660  
DB 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKQIRKILSGYVIEIDTEGLKEVINDRY 660  
QY 661 DMLNLSLRQDGTFTDFPKYNDKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNLSLRQDGTFTDFPKYNDKPLIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
DB 721 IKKILIFSCKGYEIG 735

RESULT 8  
AAM51500  
ID AAM51500 standard; protein; 735 AA.  
XX AC AAM51500;

01-FEB-2002 (first entry)  
XX  
XX Anthrax PA F427K mutant.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX  
XX Bacillus anthracis.  
XX Synthetic.

XX	SQ	Sequence 735 AA;	
XX	Query Match	99.7%; Score 3759; DB 5; Length 735;	
XX	Best Local Similarity	99.7%; Pred. No. 6.4e-248;	
XX	Matches 733; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
QY	1	EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
DB	1	EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
QY	61	QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY	120
DB	61	QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRSTAGTPVDRDN	180
DB	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRSTAGTPVDRDN	180
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTADPSPDFEKT	240
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTADPSPDFEKT	240
QY	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
DB	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
DB	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
DB	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKQDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV	480
DB	421	LNAQKQDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV	480
QY	481	LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFQNEPNGNL	540
DB	481	LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIATFQNEPNGNL	540
QY	541	QYQKQDITFDNFDOQTSQNIKNQLAELNATNIVTLDKIKLNAMNILLRDKRPFHYDR	600
DB	541	QYQKQDITFDNFDOQTSQNIKNQLAELNATNIVTLDKIKLNAMNILLRDKRPFHYDR	600
QY	601	NNIAGADESVKAEHREVINSGTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY	660
DB	601	NNIAGADESVKAEHREVINSGTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY	660
QY	661	DMLNISLRQDKTIFDFKKNYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGTSTNG	720
DB	661	DMLNISLRQDKTIFDFKKNYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGTSTNG	720
QY	721	IKKILIFSKKGYEIG 735	
DB	721	IKKILIFSKKGYEIG 735	
XX	RESULT 6		
XX	AAW51490		
ID	AAW51490	standard; protein; 735 AA.	
AC	AAW51490;		
XX	01-FEB-2002	(first entry)	
DT	01-FEB-2002	(first entry)	
XX	Anthrax PA mutant D425B.		
XX	Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;		
KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.		
XX			

OS	Bacillus anthracis.		
OS	Synthetic.		
XX	Key	Location/Qualifiers	
FX	Misc-difference 425	/note= "Wild-type Asp substituted by Glu"	
FT			
XX	W0200182788-A2.		
XX	08-NOV-2001.		
XX	04-MAY-2001; 2001WO-US014372.		
XX	04-MAY-2000; 2000US-0201800P.		
XX	(HARD ) HARVARD COLLEGE.		
XX	Collier RJ, Sellman BR;		
XX	WPI; 2002-017725/02.		
XX	Protecting humans against anthrax using mutant B groups (anthrax		
PT	protective antigens) of the pore-forming binary A-B toxin of Bacillus		
PT	anthracis.		
XX	Claim 4; Page; 77pp; English.		
XX	The invention relates to antibacterial agents comprising mutant forms of		
CC	pore-forming toxins (AAW51484 and AAW51484-AW51500), especially mutants		
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B		
CC	moiety is anthrax protective antigen (PA) and using these mutants or		
CC	compositions of them for protecting against Bacillus anthracis infections		
CC	in humans, especially as vaccines. Note: The present sequence is not		
CC	given in the specification but is derived from the Bacillus anthracis		
CC	wild-type PA protein sequence shown in figure 13 (AAW51483)		
XX	Sequence 735 AA;		
SQ	Query Match	99.5%; Score 3755; DB 5; Length 735;	
	Best Local Similarity	99.6%; Pred. No. 1.2e-247;	
	Matches 732; Conservative	1; Mismatches 2; Indels 0; Gaps 0;	
QY	1	EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
DB	1	EVKQENRLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
QY	61	QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY	120
DB	61	QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRSTAGTPVDRDN	180
DB	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNRKSRSTAGTPVDRDN	180
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTADPSPDFEKT	240
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTADPSPDFEKT	240
QY	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
DB	241	GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
DB	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
DB	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKQDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV	480
DB	421	LNAQKQDASSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYVNFENGRVVDTSNWSSEV	480



361	Db	NANIRYVNTGTAPIYVNLPTFTSLVLGKQKQTLATIKADENQLSQTLAPNNYTPSKNLP	420
421	Qy	LNAQKQDASPTPTMNVQNFLELEKTKQRLDQTVGYNIATYFNGRVRVDGTSNWSEV	480
421	Db	LNAQKDFSSPTPTMNVQNFLELEKTKQRLDQTVGYNIATYFNGRVRVDGTSNWSEV	480
481	Qy	LPOIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPMQTLKBAUKTAFGNPNGNL	540
481	Db	LPOIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPMQTLKBAUKTAFGNPNGNL	540
541	Qy	QYQGXDIITFDNFQOQTSQNIKNQLAELNATNYTVLDDIKLNAKQNWILLIRDKRPHYDR	600
541	Db	QYQGXDIITFDNFQOQTSQNIKNQLAELNATNYTVLDDIKLNAKQNWILLIRDKRPHYDR	600
601	Qy	NNIYVAGDSSVVKAEHREVINSSTEGLLNLNDKQIRKILSGYIVIEBDETLGKEVINDRY	660
601	Db	NNIYVAGDSSVVKAEHREVINSSTEGLLNLNDKQIRKILSGYIVIEBDETLGKEVINDRY	660
661	Qy	DLNLNLSLQDQGTFFDPKKYNDKPLYSINPNYKNVYAVTKENTTIINSENGDSTNG	720
661	Db	DLNLNLSLQDQGTFFDPKKYNDKPLYSINPNYKNVYAVTKENTTIINSENGDSTNG	720
721	Qy	IKKILIFSKKGYEIG 735	
721	Db	IKKILIFSKKGYEIG 735	

RESULT 3  
AAM51485  
ID AAM51485 standard; protein: 735 AA.

01-FEB-2002 (first entry)

XX DE Anthrax PA mutant K397D.

KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 KW B moxety; A-B anthrax toxin; *Bacillus anthracis*; vaccine; mutant; mutetin.  
 XX  
 OS *Bacillus anthracis*.  
 OS Synthetic.  
 OS

XX	FH	Key	Location/Qualifiers

FT Misc-difference 397

```

FT      misc-difference 397
FT      /note= "Wild-type Lys substituted by Asp"

```

XX PN WO200182788-A2.

PD 08-NOV-2001.

04-MAY-2001: 2001WO-US014372.

PR 04-MAY-2000; 2000US-0201800P.

PA (HARD ) HARVARD COLLEGE.

PI Collier RJ, Sellman BR;

DR WPI; 2002-017725/02.

PT Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of *Bacillus anthracis*.

PS Claim 4; Page; 77pp; English.

CC The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections

CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the *Bacillus anthracis*  
CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;

Query Match 99.7%; Score 3760; DB 5; Length 735;  
Best Local Similarity 99.7%; Pred. No. 5.4e-248;  
Matches 733; Conservative 0; Mismatches 2; Indels

Qy	1	EVQENRLLNBSRSSGOLLGYFSDLNFOAPMVVTSSTTGDLSIPSESELENIPSENOYF	60
Db	1	EVQENRLLNBSRSSGOLLGYFSDLNFOAPMVVTSSTTGDLSIPSESELENIPSENOYF	60
Qy	61	QSAIWSGFIKVKKSDEYTFATSDADNHVTMWDDOEVIKANSNKIKLEKGRLLQIKIQY	120
Db	61	QSAIWSGFIKVKKSDEYTFATSDADNHVTMWDDOEVIKANSNKIKLEKGRLLQIKIQY	120
Qy	121	QRENPTBKGJDFKLWYTDSONKKEVIPSDDLQIPELKQKSNSRKKGSTSAGPTVPDRDN	180
Db	121	QRENPTBKGJDFKLWYTDSONKKEVIPSDDLQIPELKQKSNSRKKGSTSAGPTVPDRDN	180
Qy	181	DGIPDLSLEVEGYTVDVKNKRTFLSPWISNTHIEKKGLTKYKSSPKWSTASDPYDFEKVT	240
Db	181	DGIPDLSLEVEGYTVDVKNKRTFLSPWISNTHIEKKGLTKYKSSPKWSTASDPYDFEKVT	240
Qy	241	GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSSTRHT	300
Db	241	GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDOSTQNTDSETRTISKNTSSTRHT	300
Qy	301	SEVHGNAEVIASFDPIDGVSAGFSNSNSTVAIDHLSLAGERTWAEWTGLN'TADTARL	360
Db	301	SEVHGNAEVIASFDPIDGVSAGFSNSNSTVAIDHLSLAGERTWAEWTGLN'TADTARL	360
Qy	361	NANIRYNTGTAPIYVLPITSLVLGKQOTLATIKADENOLSOILAPNNYPSKNLAPIA	420
Db	361	NANIRYNTGTAPIYVLPITSLVLGKQOTLATIKADENOLSOILAPNNYPSKNLAPIA	420
Qy	421	LNAQDASSTPIITMNYNOFLEKTKQLRLD'TDOVYGNIAITYNPENGRVVDTCNNSEV	480
Db	421	LNAQDASSTPIITMNYNOFLEKTKQLRLD'TDOVYGNIAITYNPENGRVVDTCNNSEV	480
Qy	481	LPQIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL	540
Db	481	LPQIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL	540
Qy	541	QYQKDLITEFDNFDQOTSQNIKNQALAEATNIYTVL'DKIKLNAKWNILIRDRKFHYDR	600
Db	541	QYQKDLITEFDNFDQOTSQNIKNQALAEATNIYTVL'DKIKLNAKWNILIRDRKFHYDR	600
Qy	601	NNIIVAGADESVVKEAREVINSSTEGILLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY	660
Db	601	NNIIVAGADESVVKEAREVINSSTEGILLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY	660
Qy	661	DMLNISLRDQKGTIDFKYKNOKLPYIISNPYKNVYATVKTENTIINPSENGD'TSTNG	720
Db	661	DMLNISLRDQKGTIDFKYKNOKLPYIISNPYKNVYATVKTENTIINPSENGD'TSTNG	720
Qy	721	IKKILIFSKXGYEIG 735	
Db	721	IKKILIFSKXGYEIG 735	

RESULT 4  
AAM51491  
ID AAM51491 standard; protein: 735 AA.

AC AAM51491;

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant D425K.

CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;

Query Match 100.0%; Score 3772; DB 5; Length 735;  
Best Local Similarity 100.0%; Pred. No. 8.2e-249;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKGSTSGPTVDPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKGSTSGPTVDPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGLKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDASSPTITWYVNCFLKLETKQLRDLTDQVYGNATYVNFGRVVDVGTGNSWSEV 480  
Db 421 LNAQKDASSPTITWYVNCFLKLETKQLRDLTDQVYGNATYVNFGRVVDVGTGNSWSEV 480  
QY 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPKNL 540  
Db 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPKNL 540  
QY 541 QYQKDIETDFNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNTILIRDKRPHYDR 600  
Db 541 QYQKDIETDFNPDQOTSONIKNQLAELNATNIYTVLDKIKLNAKNTILIRDKRPHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNLSLRQDKTFFDKYNDKPLPLYISNPYKNNVYAVTKENIINPSENGDSTNG 720  
Db 661 DMLNLSLRQDKTFFDKYNDKPLPLYISNPYKNNVYAVTKENIINPSENGDSTNG 720  
QY 721 IKKILIPSKGYEIG 735  
Db 721 IKKILIPSKGYEIG 735

RESULT 2  
ID AAM51493  
XX AAM51493 standard; protein; 735 AA.  
AC AAM51493;  
XX  
DT 01-FEB-2002 (first entry)  
XX Anthrax PA mutant K397D/D425K.  
DE

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX Bacillus anthracis.  
OS Synthetic.  
XX  
Key Location/Qualifiers  
FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"  
FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"  
XX WO200182788-A2.  
XX 08-NOV-2001.  
XX 04-MAY-2001; 2001WO-US014372.  
XX 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX Protecting humans against anthrax using mutant B groups (anthrax  
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
PT anthracis.  
XX  
PS Claim 4; Page; 77pp; English.  
XX  
CC The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;

Query Match 99.8%; Score 3766; DB 5; Length 735;  
Best Local Similarity 99.8%; Pred. No. 2.1e-248;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLGYYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKGSTSGPTVDPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKGSTSGPTVDPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 49.1208 Seconds  
(without alignments)  
4227.791 Million cell updates/sec

Title: US-09-848-909a-13

Perfect score: 3772

Sequence: 1 EVKQNRLLNSESSESSQGLL.....TSTNGIKILIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3772	100.0	735	5	AAM51495 Anthrax P
2	3766	99.8	735	5	AAM51493 Anthrax P
3	3760	99.7	735	5	AAM51485 Anthrax P
4	3759	99.7	735	5	AAM51491 Anthrax P
5	3759	99.7	735	5	AAM51492 Anthrax P
6	3755	99.5	735	5	AAM51490 Anthrax P
7	3754	99.5	735	5	AAM51487 Anthrax P
8	3754	99.5	735	5	AAM51500 Anthrax P
9	3754	99.5	735	5	AAM51489 Anthrax P
10	3753	99.5	735	2	AAR60179 Protective
11	3753	99.5	735	5	AAM51483 Anthrax P
12	3753	99.5	735	5	AAM51488 Anthrax P
13	3753	99.5	735	5	AAM51494 Anthrax P
14	3753	99.5	735	5	AAM51499 Anthrax P
15	3753	99.5	736	3	AAY56959 B. anthra
16	3753	99.5	763	3	AAY56960 B. anthra
17	3753	99.5	764	3	AAY56958 B. anthra
18	3753	99.5	764	4	AAB47306 Wild type
19	3753	99.5	857	7	ADE65872 Bacillus
20	3752	99.5	735	5	AAM51484 Anthrax P
21	3751	99.4	735	5	AAM51486 Anthrax P
22	3750	99.4	735	5	AAM518289 Bacillus
23	3750	99.4	764	6	AAB35717 Bacillus
24	3738	99.1	764	6	ABP71693 B. anthra
25	3708.5	98.3	903	2	AAR60183 PA(1-725)

#### ALIGNMENTS

##### RESULT 1

AAM51495

ID AAM51495 standard; protein; 735 AA.

XX AC AAM51495;

XX AC AAM51495;

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant K397D/D425K/F427A.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX Bacillus anthracis.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 397

FT Misc-difference /note= "Wild-type Lys substituted by Asp"

FT Misc-difference 425

FT Misc-difference /note= "Wild-type Asp substituted by Lys"

FT Misc-difference 427

FT Misc-difference /note= "Wild-type Phe substituted by Ala"

XX WO200182786-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD ) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

XX Protecting humans against anthrax using mutant B groups (anthrax

PT protective antigens) of the pore-forming binary A-B toxin of Bacillus

PT anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of

XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants

XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B

XX moiety is anthrax protective antigen (PA) and using these mutants or

XX CC



Db 1395 NKWHNFINSYFDLNENEKKKKINIF 1419

Search completed: May 3, 2004, 19:40:03  
Job time : 38.6261 secs

[illegible]

RESULT 15

Q81HP9 ID Q81HP9 PRELIMINARY; PRT; 2940 AA.  
AC Q81HP9;  
AC Q81HP9;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, last annotation update)  
DE Hypothetical protein.  
DE PF11\_0490.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxId=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3D7;  
RX MEDLINE=22455705; PubMed=12369864;  
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
RA Carlton J.M., Patn A., Nelson K.E., Bowman S., Faulsen I.F., James K.,  
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,  
RA Perten M., Allen J., Sellent J., Haft D., Mather M.W., Vaidya A.B.,  
RA Carter M.,

QY	682	NDKLPPLIISPNYKNVYAVTKENTIIINPSENGDTS	71.7
Db	4416	SNKPL- KLNINNYQIK--INKSHTII--SKNGEWS	44.45
RESULT 13			
QBIHY4	QBIHY4	PRELIMINARY;	PRT; 2849 AA.
AC	QBIHY4		
DT	01-MAR-2003	(TREMELrel. 23, Created)	
DT	01-MAR-2003	(TREMELrel. 23, Last sequence update)	
DT	01-MAR-2003	(TREMELrel. 23, Last annotation update)	
DE		Hypothetical protein.	
GN	Pf11_0392.		
OS	Plasmodium falciparum (isolate 3D7).		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
NCBI	Taxid=36329;		
TRN	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=3D7;		
RC	MEDLINE=22355705; PubMed=12368864;		
EX	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,		
EX	Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,		
EX	Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,		
EX	Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,		
EX	Pertea M., Allen J., Selengut J.J., Haft D., Mather M.W., Vaidya A.B.,		
EX	Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,		
EX	McFadden G.I., Cummings L.M., Subramanian G.M., Muggall C.,		
EX	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,		
EX	Fraser C.M., Barrell B.		
EX	"Genome sequence of the human malaria parasite Plasmodium		
EX	falciparum."		
EX	Nature 419:498-511(2002).		
DR	EMBL; A5014841; AAN35975.1; -.		
DR	Hypothetical protein.		
SW	SEQUENCE 2849 AA; 335916 MW; B5515D173D96B813 CRC64;		
Query Match 5.3%; Score 201.5; DB 5; Length 2849;			
Best Local Similarity 20.6%; Pred. No. 0.67;			
Matches 170; Conservative 134; Mismatches 333; Indels 189; Gaps 41;			
QY	3	KOENRLLNSESSESSQGLGY-YPSDLNFO--APMVTSTTGDLSTP--SSELENTPSEN	57
Db	1862	KHNNYTFNKEENNLKIFKYNNNNNINEXDDIPKSIQSSFIEDNIAAYNKGNNKKLN	1921
QY	58	Y-FQSAIWGGFTKVKSDPEYTFATGADNHVTWVDD-----QEVINKASNSK----	105
Db	1922	EDIIQERFSHNNTKTYEMNKNSYDSCDNIIVINDELNDSTQTKELNKSNGKAEAW	1981
QY	106	-IRLEKGRLY-QIKIQYQRENPEKGLDPKLYWTDQNKKEVTSQNLQPELKQSSNS	163
Db	1982	IIDIKNNETYPYIKDKGKKNEDK--KQKMYNKNQDNKNIKMGSSN--KSMKQNKNS	2036
QY	164	RKGRSTAGTVPDRDNDGIPDSLEVEGYTVDYVKNKRTFLSPWISNHEKKGLTKYKSGP	223
Db	2037	NQMKH-----IPUSVNNKGY-----NKSSNKKYENNIN-----KNNK	2069
QY	224	EKWSTASDPYDF-EKVTGRI-DKNVSPPEARHPLVAAYPIVHDM-----NIIIS	272
Db	2070	DKLNLVNSISKLVQSKIKQELSNKNISKD-----ILNFEITKIKKSKKETQNTWHTNN	2124
QY	273	KNEDQSTQNTDSETRISKNSTSRTHTEV-HGNAEVEASFPDIGGSVAG-FSNSNS	330
Db	2125	KNDNNNDNNINFERINDINNVNLNKNKYVDNINI-----DEGNYINGYFIYAPDS	2179
QY	331	TVAIHDSLSLACERTWAETWGLNTA-----DTARLNA-----NIRYVNTGTAPYNNL	378
Db	2180	TI-----YTWVNIDITWYMKYFDSKKNINHVPLNNVPLN-----NDL	2221
QY	379	PTTSLVLGKQNTIATDAENQLSQ-----ILAPNYYPSKQLAPIALNKKQFSPTITW	434
Db	2222	LINNVLNQNMMNLENNTNTIGSQVPFVTCDFY--ANNIKSIYLD-----PNLQ	2271

DR EMBL; AE014841; AAN35955.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 3468 AA; 411913 MW; 5067D62999BA7B1 CRC64;  
  
Query Match  
Best Local Similarity 18.9%; Pred. No. 0.36;  
Matches 158; Conservative 130; Mismatches 276; Indels 271; Gaps 37;  
  
QY 55 SENCYFQSAWSGFIKVKSDYETFTATSADNHYTMVMDQEV---INKASNNKIRLEK 111  
DB 160 SENEY-----TKNDCKNNSTNSGY-----NKEIEFLKNIAKEHSLKIEKN 202  
QY 112 RLVOIK-----IQVQRENPEKGLDFK---LYM---TDSON-----KKEVTSOHL 151  
DB 203 INLLKCDSDNNDLBPKKNNVHLNNDVASEDLFHDNRVNNQNLKIDIKYIHHDDMI 262  
QY 152 QLPKLGKSSNRKRSSTAGTPVDRDNDGIPDSLEVEGYTVDV--KNKRTTLPSPWISNI 210  
DB 263 MI-----ESNDRNDKNCFNTPNYERHIIVDKTKNGKNNNNINSNNFDS--ISNI 315  
QY 211 HEKGLTKYKSSPEKSTASDP---YSDFE-KVTGRIDKQVSPPEARHPLVAAYPIVHV- 264  
DB 316 NKK-----ISYPINMYNSSEDKTLNQLNLDMSILSDSLKNAAYSPYSLS 360  
QY 265 ---DMENIILSKNDQSTQTD--SETRTSKNTSTRTSTSEVHGNAEVAH--SPFDIG 317  
DB 361 EKNNNGNNSHKNNNYRDVIDEDINISKFSINLHSDSENNDFNSLCSP----416  
QY 318 GVSAGFSNSSTVAIDHSLSLAGERTWAEITMGL-----NTADTARLANIRY---- 366  
DB 417 ---NSSKCVKDEMITQFVGNRYKYKVTMDAMNDYMKNSIEQHSINMESNHFKNQ 468  
QY 367 -VNTGTAPIYVLPVTSVLGKQKQ-----TLATIDADE--NQLSILAPNN--- 409  
DB 469 KINSKGEDANNLNSHILNNKIGIQVNSLSDYLSIKKQESWNNNSNNALNNIN 528  
QY 410 -----YPSKNLAPI-----ALNAQKFSSTP--431  
DB 529 TNNNGSNNKNSNVYKTSQYVYVNPNDNINNMWHLNLSYMKNSNNLNTSNGFKKIPKN 588  
QY 432 -----ITMNVNQLFLEKTKQRLDTPQVYGNIAVTYFNGRVVDGTGNSWSEVLPOQ 485  
DB 589 KNIISNIDFDNNIFKSYIKENVVKQNSINHOVEKQNYTNDSEINKN-----NNIENNTQ 644  
QY 486 ETTAR-----IFFNGKDLNLVERRIAAYVPSDPLETTK-----518  
DB 645 NTTCNFINTNDIINKKIKIYKIKIDSISLLNDLSLNKLESINSIMDRYTKNNYEEK 704  
QY 519 --PDMTLKEAL-----KIAFGNEPQNLQYQKQDITEFDNFDQOTS 559  
DB 705 FLDDVILDDSIATSNELOHSNYTTTTHIFDNNNNNNNNINQKEDLFQNDYN---KET 761  
QY 560 QNKNQALAEINATYVLDKIKLNKAKNIIIRDKREPHYDRNNIAYGADESVVKEAREV 619  
DB 762 YNNIMISENNAENLF-----KISYSCDLVLGKNEIILD--RVV 800  
QY 620 INSSTEGLLNIDKIRKILSGYIETEDTEGLKE-----VINDRYDMLMISS--L 668  
DB 801 ENSKTE-----QVEYQNKEDIKLYHKDDNIIIDNNYNDVNIKKDCHL 844  
QY 669 RODGKTITDFKYN--KLPL-----YISNPNKYV--VYAVTKNTIINPSEN 713  
DB 845 KMDNQDNKINKSQSKKKNPINNECNVHNKFNKINEIDHPKEENTLAYSCKN 899

RESULT 12  
Q9PQ08  
ID Q9PQ08 PRELIMINARY; PRT; 4688 AA.  
AC Q9PQ08;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein UU482.

GN UU482.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Serovar 3;  
RX MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RA Caselli G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum";  
RL Nature 407:757-762(2000).  
DR EMBL; AE002145; AAF30894.1; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 4688 AA; 534880 MW; B53ABFAFEE1997E CRC64;  
  
Query Match 5.4%; Score 202; DB 16; Length 4688;  
Best Local Similarity 21.3%; Pred. No. 1.2;  
Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;  
  
QY 19 LLGYFSDLNFQAPMVYTSSTGDLSPSSSELENIPSENOYFOSAIWGSFKVKVSEY 78  
DB 3699 LVDVYVLD-NIHQNIIDETRIKIFKDHV-SKEIEINPGVTWISKHGNWKSPTDTTANFEFK 3756  
QY 79 FATSADNHYTMVMDQVINKASNNKIRLEKGLYQIKYQORENPTKGLDFKLYWTD 138  
DB 3757 IETQ-----DNDVNLNIDATVKDHEHNNIKOKIVRIKEN-----ND 3795  
QY 139 SONKEVISSNLQPLKQKSSN---SRKRSSTAGTPVDRD-----179  
DB 3796 WLIQNI---DNLN-PETKYKLENIELSKPLKTHNLVSINDKENISLITETGNPVLKV 3851  
QY 180 ---NDGTPDSLEVEGYTVDVKNKRTTLPSPWISNIHEKGLTKYKSSPEKSTASDYS 235  
DB 3852 IOTQNTDITNDTQNTVTLGVSNSK-YNGROIKVYKNNNNVYESS---LITLQKQND 3907  
QY 236 FEKVTGRIDKQVSPPEARHPLVAAYPIVHVDMEINILSKN-EDQSTQNTDSET-RTISKNT 293  
DB 3908 YQLLLSNLNSN-----REYRFEKIEINHSINTNPFEDLEKLGVSNTFITQKNT 3957  
QY 294 STSRTHTS-EVHGNAEVHASP-----FDIGSVSAGRS-----NSNS--STVAID 335  
DB 3958 TVQWNSDSATTIVGRVGNFNFKIKSEDKILENNOQVAVFAPKETIRDTWLTQYRPLK 4017  
QY 336 HSLSLAGERTWAEITMGLNT---ADTARLANIRVYVNTGTAPIYVLPVTSVLGKNQTLA 392  
DB 4018 DVTSDFKEGTWAHDLNSNVNFEETTYKLVKIQFVNKETKAKNNINSENNVILDNTNSI 4077  
QY 393 -----TIDADENQLSILAPNYYPSKQLAPIALNAQK--FSSTPTIMNY--NQFLE 443  
DB 4078 NSNYEFTTKVGDHKLINITSSNNVTNSQTINFTLSGVKYSWVGKIKLSYKSDTSESI 4137  
QY 444 KTKQLRLDTPQVYGNIAVTYFNGR-----VRVDTGNSWSEVLPOIQTETPARIIFNGKD 497  
DB 4138 HTNEVLIESNKTQYNILLNLLNKKRNTYTLIDVLDIDNNVSDFPKEGNLTNSFITRTSA 4197  
QY 498 LNLVERRIAAVNPSDPLETT-----KPDWTLKEA-----LKIAFGNEPQNLQ 541  
DB 4198 INVLNIEISNRASTNLKSTIIKINLNDPDNVLKDQQAATTIVYGNKQAMGFIIVSGNIK 4257  
QY 542 YQKQDITEFDNFDQOTSQNIK-NOLAEINATYVLDKIKLNKAKNIIIRDKREHYD- 599  
DB 4258 YLTATVLDLNFN-DKVNIVNISFNKPSIAAEN-----IGDKSNII-----YRDS 4304  
QY 600 -----RNNIAYGA---DESIVKEAREVINSSTEGLLNIDKIRKILSGYIETEDTE 650  
DB 4305 IPKLEINNDIIVNGPINKKEIVVKANQK--NNIDVDLGLQINPKIAHLNR-FIAKFKSTN 4361  
QY 651 GLKEVIND--RYDMLNITSSL-RDQKTFIDF-----KKY-----681  
DB 4362 -----NDIETNVINGSSLVNNDGKTSIRFTLNNLKANKLYSLVDVYVYVNNNSNTIVE 4415

Db	1	MESLGNNITNALDRIKUNAKNNILVRDP- YHYDNNGNIVGVDDSYLKNAYKQILNWSSD	59
Qy	626	GLLNLNDKDIRKILSGYIEIEDTE----- GLKEVINDRYDMLMISSLRDQ	672
Db	60	GVSNLNDEDVNAQSGYMLQIKKPSNHLTNSP VITITLAGKDSGVGELYRVL:-	113
Qy	673	KTFIDFKKNDKLPYISNPKNYKVVATKEN- TIINPSENGDSTNGIKKILIFSCKG	731
Db	114	TGFLDFNKFEDNWSLV-DPGDDVYVAVATK EDFNAVTRDENGRIA-NKLKNTLVLSGKI	171
Qy	732	YEI 734	
Db	172	KEI 174	
RESULT 10			
Qy	Q8RGK2	PRELIMINARY; PRT; 1881 AA.	
AC	Q8RGK2;		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Hemolysin.		
GN	FN0291.		
OS	Fusobacterium nucleatum (subsp. nucleatum).		
OC	Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;		
OC	Fusobacterium.		
OX	NCBI_TaxID=76856;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 25386;		
RX	MEDLINE=21886394; PubMed=11889109;		
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,		
RA	Bhattacharyya A., Bartman A., Gardner W., Greshkin G., Zhu L.,		
RA	Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,		
RA	Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,		
RA	Fonstein M., Kyripides N., Overbeek R.,		
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium		
RT	nucleatum strain ATCC 25386.";		
RL	J. Bacteriol. 184:2005-2018(2002).		
DR	EMBL; AEO10541; AAL94497.1; -		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0008233; F:peptidase activity; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR008638; Haemagg_act.		
DR	InterPro; IPR000508; Peptidase_S26.		
DR	InterPro; IPR001452; SH3.		
DR	Pfam; PF05860; Haemagg_act; 1.		
DR	Pfam; PF00018; SH3; 1.		
DR	PROSITE; PS00761; SPASE_I_3; 1.		
KW	Complete proteome.		
Qy	SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;		
Query Match			
Best Local Similarity 6.0%; Score 226.5; DB 16; Length 1881;			
Matches 176; Conservative 131; Mismatches 302; Indels 233; Gaps 373			
Qy	30	QAPMVVTS-----STTG----- -----DLGIPSSLE---NIPSENOY-FOSAIWSGPI	69
Db	246	QAPVATKGDVVISGKGYLKDTQAKEDIKISTETIGSKLLAEWAINISGKTSNSG	305
Qy	70	KVKSDSEYTFATSDAHVMTVMVDDQEVIN----- -----KASNSKIELEKGR	112
Db	306	QIPANNNTINGNVDSNLIFTNKKDITISGNLKNNGVSSNLSNLNKEIENSNNKVVVEE-K	364
Qy	113	LYQIKI-----CYORENPTEKGLDFKLYWT-- -DSQNKKEVISDNLQLPELKOKSS	161
Db	365	LSSTKTLNLCNLSAKEIEKTNIFNSGKLPFSKNIITAKDFKNGEV-SSENLTTTNLSNWK	423
Qy	162	-----NSKKKSTSGAGTVPRDNDG----- -----	182
Db	424	INVKENINGSIVNKTNAEITSKNLNTNLNDRNGITIIINNVSQVIANNGKLLVGNNTIN	483

```
QY 226 WSTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEQDQSTQNTDSE 285
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 231 PCTANDPYDPEKVSQGDPSVMSVARDPMISAYPIVGVQMERLVVSKSE-----TITGDS 286
QY 286 TRTISKNTSTGRTTSEVHGNAEVSASFDDIGG-----SVSAGFSNSNSTVAIDHSL 338
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 287 TKSMKSTSHSTNINTV--GRAVSGSLQAGGPPVFSMSASANYSHTWQNTSTVDDT- 343
QY 339 SLAGERTWAETWGLNTADTARLANIRVNTGTAPIVNVLTPTTSLVLGKQNTLATIDAE 398
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 344 --TGE-SFSQGLSINTGSAVINPINRINYNTGTAPVNVVTTTIVIDK-OSVATIKGQE 399
QY 399 NQLSQILAPNNYPYPSKNLAPIALNAQKFFSTPTIMYVNVQFLEKTKQLRLDQVYGN 458
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 400 SLIGDYNLPGGTYPIIGEPFPAALNTMDOFSSRLPIVINYQLKSIDNGGTWMLSTSQFTGN 459
QY 459 IATYVFNENGRVAVDTGSNWSVLPQIQETTARIL--FNGKDLNVERRIAAVNPSPDLET 516
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 460 FAKYN-SGNLVTD--GNWNGPYLGTIKTGTASLTLSFSGQTQVA---VWAPNPSDPEDK 514
QY 517 TKPDMTLKALKIAPG--FNEPNGNLQYQCKDIT--EPDFNFDQQTSONIKQLAELNATN 573
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 515 T-PKUTLQALVKALFKKQKGFYFHGLEISKNEKIQVFLDSNTNDFENQLKNTADKD 573
QY 574 IYTVLDKIKLNARQMILIRDRPHYDRNNIAV 605
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 574 IMHCI--IKRN--MMILVKVITFKENISSINI 601

RESULT 8
Q844J8 PRELIMINARY; PRT; 775 AA.
ID Q844J8
AC Q844J8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ViplAC.
GN ViplAC.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Pang Y.;
RT "Cloning of vipA(C) and vip2A(C) from Bacillus thuringiensis HD201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY245547; AAC86514.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; BINARY_TOXINB.
DR PRINTS; PR01391; BINARY_TOXINB.
SQ SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;

Query Match 19.1%; Score 721; DB 2; Length 775;
Best Local Similarity 29.0%; Pred. No. 5.2e-26;
Matches 223; Conservative 133; Mismatches 294; Indels 120; Gaps 30;

QY 4 QENRLNESSESSQGLLYPDLNFCAPMVVTSSTGDLSPSSELEN--IPSENQYFQ 61
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 42 QKQK--QKEMDRKGLLYGFKGDF-SNLTWPAFTRDNTLIYDQQTANKLDRKQQQYQ 97
QY 62 SATWSGFIKVKSDVEYFATSDAHVMTWVDQEVNKASNKRLEKGRLYQIKIQY 121
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 98 SIRWIGLQSKGKDFTNLSEBEQALIEDGKIISNKGKQVHLEKLVPIKIEYQ 157
QY 122 RENPTEKGLD-----FKLWVTSQNKKEVISSDNLQLPELKQKSN----- 162
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 158 SD--TKFNIDSKTFPKFKLIDSKQSQVQKRDLENFENFKKESREFLAKASKTNPFM 215
QY 163 SRKRSSTAGTVPDRNDGIPDSLEVEGVTVDVKNKRTFLSPWISNIHEKGLTKYKS 222
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```

RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 10748;
RX MEDLINE=94041637; PubMed=8225592;
RA Ferelle S., Gilbert M., Boquet P., Popoff M.R.;
RT "Characterization of Clostridium perfringens iota toxin genes and
RL expression in Escherichia coli.";
RL Infect. Immun. 61:5147-5156(1993).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=NCIB 10748;
RC STRAIN=NCIB 10748;
RA Popoff M.R.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X73562; CAA51960.1; -
DR FJ; 140862; I40862.
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; F:toxin activity; IEA.
DR Pfam; PF03495; B:binary toxin; IEA.
DR PRINTS; PR01391; BINARYTOXINB.
KW SIGNAL.
FT SIGNAL.
FT CHAIN.
FT CHAIN.
SQ SEQUENCE 875 AA; 98468 MW; C9AE092CD3818921 CRC64;

Query Match
Best Local Similarity 31.1%; Score 871; DB 2; Length 875;
Matches 252; Conservative 131; Mismatches 278; Indels 150; Gaps 31;

QY 1 EVKQENRLNESSESSGGLGYFSDLPNFAPMVWTSSTTGGDLSIPSSSELENIPSE-NQY 59
DB 35 DTNKEEINTENLTSSNGLMGYFDEHFKDLELAPIKNGDLKFEKKVKKLLTSS 94
QY 60 FQSAIMSGPIKVKSDYEYFATSADNHVTMWVDDQEVINKASNKIRLEKRLYQIKI 119
DB 95 IKSIRWTGRIIPSEDEYILSTDR-NDVLQINAKGDIK---TLKVNMMKGQAVNRIE 150
QY 120 YQEN-----PTEKGLDKLYWDSQNKVEISSDNLQLPKQKSSNRKRKSTS 170
DB 151 IQDKNIGSDINLQVSP-----KLYW-ELNGKNTVPEENLFFRYSKIDEND----- 195
QY 171 AGPTVP-----DRDNDGIPDGLSEVEGYTVDVKNKRTFLSPWISNIH 211
DB 196 --PFIFNNFFDVRPFSAWEEDELDTDNDNIPDAVEKNGYI-----KDSIAVKNDQSA 249
QY 212 EKGLTKYSSPEKWTASDPSDFEKTGRIDKQVSPPEARPLVAAYPIVHVDMENIL 271
DB 250 E-QGYKKYSSVLESNTAGDPTDYQKASGSDKAIKLEARDPLVAAYPVVGVMENLII 308
QY 272 SKNEDOSTQNTDSERTIGKNTSTSTHTSEVHGNAEVHASFDDIGGSVSAFSSNSGST 331
DB 309 STNEHASS-----DQKTVSRATNSKTANTV-----GVISAGYQNGFTGN 351
QY 332 VAIDHS-----LSLAGERTWAETMGLNTADTARLNANIVNTGTAPIYVNLPTSL 383
DB 352 ITTSYGHNTDSTAVQDSNGESWNTGLSINKGESAVINANRYNTGTAPMYKVTPTNL 411
QY 384 VLKKNOTLATIDENQLSQILAPNNYPSKNLAPLNAOKKFSSTPTWNNQFLE 443
DB 412 VL-DGELTATIQADQIGNUSPNETPKGLSPALMTQOFNARLPIINDYQKLLD 470
QY 444 KTKQLDLTDQVYGNATYFENGVRVDTGNSWSEVLFPQIQTETARIIFNGKOLNVER 503
DB 471 SGQIKLETTQVSGNYGYN-SQGI-I-TEGNSWSNYISQIDSVSASIID-TGSOTFER 527
QY 504 RIAAVNPSPLETKPDMILKALAFGNFENGNYQY-QGKDIIE--FDNFQOQTSQ 560
DB 528 RVAAKSGQNPEDKT-PEITIGEAIKKAFSATK-NGELLYFNGIPIDESCVLEIFDNTSE 585
QY 561 NIKNQLAELNATNIYVLOKILNNAKONILIIDKRF--HYDR-NNIAVGADESIVVKEAR 617

```

```

DB 586 IIEQLKYLDKKIYVNV---KLRGMNLIKVPSTYFTNFDEYNNFP--ASWSNIDTKNQ 639
QY 618 EVINSSTEGI-----LLNIDKDIRKILSGY-----IVEIEDTEGLKEVIND 658
DB 640 DGLQSVANKUSGETKIIPMSKLRPKRYVFSGYKDPSTNSITVNIKSKEOKTDYLPV 699
QY 659 RYDMLNIS-----SLRQDQKTFIDFKKYND--KLPLYISNPYKV----- 696
DB 700 EKDYTFSEYFETTKGDSSDIEITLTSSGVIPLDNLISITELNSTPEILKEPKVPSDOE 759
QY 697 -----NVYATVKENTIIINPSENGDTSTNGI 721
DB 760 ILDAHKKYADIKLDT-----NTGNTYIDGI 785

RESULT 7
ID 086171 PRELIMINARY; PRT; 721 AA.
AC 086171
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE C2 toxin (Component-II).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=98323874; PubMed=9659689;
RA Kimura K., Kubota T., Ohishi I., Isogai H., Isogai E., Fujii N.;
RT "The gene for component-II of botulinum C2 toxin.";
RL vet. Microbiol. 62:27-34(1998).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=type C;
RX MEDLINE=96184657; PubMed=8645309;
RA Fujii N., Kubota T., Shirakawa S., Kimura K., Ohishi I., Moriishi K.,
RA Isogai E., Isogai H.;
RT "Characterization of component-I gene of botulinum C2 toxin and PCR
RT detection of its gene in clostridial species.";
RL Biochem. Biophys. Res. Commun. 220:353-359(1996).
DR EMBL; D88982; BAA32537.1; -
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 721 AA; 80515 MW; 44C8153AC749D5F2 CRC64;

Query Match
Best Local Similarity 35.3%; Score 813; DB 2; Length 721;
Matches 223; Conservative 118; Mismatches 227; Indels 64; Gaps 25;

QY 3 KQENRLNESSESSGGLGYFSDLPNFAPMVWTSSTTGGDLSIPSSSELENIPSENYF 60
DB 5 KFENSVKSNKNYFTINGLMGYFEN-DFENLAIISPTLDGNTFSKEDINSILG-NKII 62
QY 61 QSAIWSGFIKVKSDYEYFATSADN-HVTMWVDDQEVINKASN-SNKIRLEKRLYQIKI 118
DB 63 KSAEWIGLIRPSITGEYILSTNSPCKVEL---NGEIFNLSTNTVNIQGNVYDIRI 119
QY 119 -QYQRENPTEKGLD-PKLYWTDQNKKEVISSDNLQLPKQKSSNRKRKSTSAGPTV- 175
DB 120 EQLMSENQLLNKNGEIKLYWETSDIKEIIPSEVL-----LKPNSVNTNEKSKTIPNLT 175
QY 176 -----PDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEK 225
DB 176 SNAKLKANANRDTDRDGPDEWEINGYTVNMQAVAWDDKFAAN-----GYKKYVSNPFK 230

```

```
Db 321 ----DOGKTVSRNTTNSKTDTANTAGVAINIAYQNGF---TSGITTYNSHTTENSTAVQNS 373
Qy 338 LSIAGERTWAETMGLNTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIDAD 397
Db 374 ---NGE-SWNTSLSINKGESAYINANRYVNTGTAPYKVTPTTNLVL-DGDTLTATIKAQ 428
Qy 398 ENCLSQILAPNNYPKSNLAPIALNAQKFSSTPTIMYNOFLELEKTKQLRLDQVYVG 457
Db 429 DNOIGNNLSPNETYPKGLSPLALNTMDQFSRLPIPYDQKJLQAGKQIKLSTTQVSG 488
Qy 458 NIATYFNGRVRVDT-GSNMSEVLPOIQTETARIIFN-GKDLNLVERRIAAVNPSPLE 515
Db 489 N---YGIKNSQGIITEGNSWSYISQIDLSASIIIDTQSD--VFERRVTAKDGSNPED 543
Qy 516 TTPEDMTLKEALKIARFNEPNGNLOYOQKDKITE--FDNFQOQTSQNIKNQIASLNATN 573
Db 544 KT-PVITIGAEIKAFKATNGKILYFNGMPIDESCVELIFDGTANLKERLNALNDKK 602
Qy 574 IYTVLDKIKLANAKNLIIRD----- 593
Db 603 IYNV---QLERGKILIKTSTYFNNFDGYNPPSSWSNVDSNNQDGLQNAANKLSGETK 658
Qy 594 -----KRPHY----- 598
Db 659 IYPMKSLNPKRYVFSGYLKNSTSNPIVNIKAKEXQTYNLVSNENDYKFSYBFTIG 718
Qy 599 -DRNN-----IAGADESVVKEAHEV-----I 620
Db 719 RDASNIEITLSSGTIFLDMLSITELNSTPEILKEPDIKVPSDQBI-DHKKYADLSF 777
Qy 621 NSSTEGLLN-----IDKDIRKILSGVIVIE-DTEGLKEVINDRYDMNLNISLROD 671
Db 778 NQSTANYLDGLYFEPTQNKVELDYIQKYVBEATLEYSGFKDIGTKKELANYTGDSNQ 837
Qy 672 GKT-FIDFKYNDKPLUYISNPY---KVVYAVTKENTIINPSENGDTSTNGIKKILI 726
Db 838 PKTYNVNFRSY-----FSGENVMPYKLIIRIATIPEN-----KELLV 875
Qy 727 FS 728
Db 876 LS 877

RESULT 5
Q9KH41 ID Q9KH41 PRELIMINARY; PRT; 876 AA.
AC Q9KH41;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CdB.
GN CDB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1496;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 20309;
RA Chang S.Y., Song K.P.;
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain
RL submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF271719; AAF81761.1; -.
DR HSSP; P13423; IACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB; I.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;
```

```
Query Match 23.2%; Score 875.5; DB 2; Length 876;
Best Local Similarity 31.4%; Pred. No. 3.4e-33;
Matches 250; Conservative 140; Mismatches 280; Indels 125; Gaps 30;
Qy 4 QENRLNSESQGLGYYFSDLNFPQAPMVVTSSTTGLSIPSSLEN-IPSENOYFOS 62
Db 40 KKKEIVNEDILPNNGLMGYYFTDEHFKDLKMAPIKDGNLKPEEKVLDLDDKSDVKS 99
Qy 63 AIWSGFTKVKKSEDEYTPATSDAHNVHTWVDDQVINKASNSKIRLEKGLVQIKYQOR 122
Db 100 IRWTGRIPSKDGEYTLSTDRDD-VLMQVNTSTI---SNTLVNKKMKKGEYKVRLELQD 155
Qy 123 EN--PTEKGLDFKLXMTDSQNKKEVTSNDLQLPKQKSSNRKRSSTAGTVP--- 176
Db 156 KNLGSIDNLSPNLYW-ELDGMKKIPEENLFLRDY-----SNIEKD-----DFIFPNNF 205
Qy 177 -----DRNDGIPDSLEVEGTVVKNKRPFLSPWISNIHKKGLTKYKSS 222
Db 206 FDPKLMDEWEDLDTNDNIPDSYBERNGYTI----KDLIAVKEWDSFAB-OGYKKYVSN 260
Qy 223 PEKWTASDPSYDPEKVTGRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNT 282
Db 261 YLENTAGDDPYTYEKASGDFKAIKTEARDPLVAAVPIVGVGMEKLIISTNEHAST--- 317
Qy 283 DSETRTISKNTSTSRHTSEVHGNAEVHASFIDIGSVSAGFSNSNSSTVAIDHS----- 337
Db 318 -DOGKTVSRATTSKTESNTA-----GVSNVGYQNGFTANVTYNSHTTDN 363
Qy 338 ---LSLAGEATWETMGLNTADTARLANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATI 394
Db 364 STAVQDNGESWNTGLSINKGESAYINANRYVNTGTAPYKVTPTTNLVL-DGDTLSTI 422
Qy 395 DADENQSQILAPNNYPKSNLAPIALNAQKFSSTPTIMYNOFLELEKTKQLRLDQD 454
Db 423 KAQENQIGNNLSPGDTYPPKGLSPLALNTMDQFSRLPIPYDQKJLQAGKQIKLETTQ 482
Qy 455 VYGNIAIYFNGRVRVDTGSMSEVLPOIQTETARIIFNGKDLNLVERIAAVNPSPDL 514
Db 483 VSGNFGTKN-SSGQI-VTEGNSWSDYISQIDLSASIIIDTEN-ESYERVTAKNLQDPE 539
Qy 515 ETTKPDMTLKEALKIARFNEPNGNLOYOQKDKITE--FDNFQOQTSQNIKNQIAELNAT 572
Db 540 DKT-PELTIGEAIEKAFKATKDGLLYFNDIPIDESCVELIFDNTANKIKSLKLSLSDK 598
Qy 573 NIYTVLDKIKLANAKNLIIRDKRFHY---DRNNIAGADESVVKEAHEVINSTEGL-- 627
Db 599 KIYNV---KLERGMNLIKTPYTFNFDYNNYP--STWSNVNTTKOGLQGSANKLNG 652
Qy 628 -----LLAIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656
Db 653 ETKIKIPMSELKPYKRVVFSGYSKDPLTSNSIIVKIKAKEEKTDLVLPQGYTKFSYEFE 712
Qy 657 NDRYDMLNIS-SLRQDKTFIDFKKYND--KLPLYISNPYK-----VNVYAVTKENT 706
Db 713 TTEKSSNIEITLIGSGTGYLDNLISITELNSTPEILDEPEVKIPTDQEIIDAHKIYFADL 772
Qy 707 IINPSENGDTSTNGI 721
Db 773 NFNFS-TGNTYINGM 786

RESULT 6
Q46221 ID Q46221 PRELIMINARY; PRT; 875 AA.
AC Q46221;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Iota toxin component Ib precursor.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1502;
```



```
AC Q32739;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE ADP-ribosyltransferase.
GN CDTB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1496;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD196;
RX MEDLINE=97230316; PubMed=9119480;
RA Perelle S., Gilbert M., Bourlioux P., Corthier G., Popoff M.R.;
RT "production of a complete binary toxin (actin-specific ADP-
RT ribosyltransferase) by Clostridium difficile CD196."
RL Infect. Immun. 65:1402-1407(1997).
DR EMBL; L76081; AAB67305.1; -.
DR HSSP; P13423; 1ACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxins.
DR Pfam; PF03495; Binary toxin; I.
DR PRINTS; PR01391; BINARYTOXINB.
KW Transferase.
SQ SEQUENCE 876 AA; 98797 MW; 25E06E2D45CE2B3B CRC64;

Query Match 23.3%; Score 877.5; DB 2; Length 876;
Best Local Similarity 31.8%; Pred. No. 2.7e-33; Indels 125; Gaps 30;
Matches 251; Conservative 139; Mismatches 280;

QY 4 QENRLNESSSOGLLYGYSFSLNFCAPMVVTSSTGDLSPSSELEN-IPSENQYFQS 62
DB 40 KKEIVNEDILPNNGLMGVYFSDHFKDLKMAPIKDGNLKPEEKVDKLLDKSDVKS 99
QY 63 AINSGFTKVKSDYEYFATSADNHVTMWYDDQEVINKASNKIRLEKGRLYQIKIQOR 122
DB 100 IRWTGRIIPSKDGEYTLSTDRDD-VLMQVNTSTI---SNTLKVNMKGKRYKVAIELOD 155
QY 123 EN--PTEKGLDFLKYTDSQNKKEVISSDNLOLPELKQSSNSRKRSTSAQTPVP---- 176
DB 156 KNLGSDINLSSPNLYW-ELDGMKKIIPENLFLRDY----SNIEKD----DFFIPNNFV 205
QY 177 -----DRNDGIPDSLEVEGYTVDVKNKRTFLSPFWSNHEKKGILTKYSS 222
DB 206 FDPKLMSDWEDELDTDNDNIPDSYERNGYTI----KDLIAVKWEDSPAE-QYKKYVSN 260
QY 223 PEKWTASDPSYDFEKTGRIDKNGSPEARHPLVAAYPIVHVDMENILSKNEDOSTQNT 282
DB 261 YLSNTAGDPTDYEKASGSFKAITEARDPLVAAYPIVGVGMKLLISTNEHAST--- 317
QY 283 DSETRISKNTSTSRTHTEVHGNAEVHASFDDIGSVSAGFSNNSSTVAIDHS----- 337
DB 318 -DQKTVSRATNTSKTESNTA-----GVSVMVQNGPTANVTNYSHTTDN 363
QY 338 ---LSLAGETWAEATWGLTADTARLNANIRVYNTGTAPIYVLPPTSILVGNKNTLATI 394
DB 364 STAVQDSGBSWNTGLSINKGSAYINAVRYNTGTAPWKVPTPINVLV-DGDTLSTI 422
QY 395 DADENQLSQIALFNYPYKSLAPIALNAQKFSSTPITMNYNQFLELEKTKQLRLDTQ 454
DB 423 KAQENQIGNLSPDGYPKGLSPLALNTWDQPSRLIPINYDQLKLDAGQIKLETQ 482
QY 455 VYGNATYFENGVRVDTGSNWSVLPOIQTETABIIIFNGKDLNLVERRIAAVNPSP 514
DB 483 VSGNFGFKN-SSGOI-VTEGNSNSDYISDTSISIIIDTEN-ESYERRVAKNLQDDE 539
QY 515 ETTKPMTLKEALKAFGFPNEPKNLQYQKQITE--FDNFQDQTSQNIKNQLAEALNAT 572
DB 540 DKT-PELTICEAKAFATKDGLLYFNDIPIDESCVELLFPDNTANKIKDSLKLSDK 598
```

```
QY 573 NIYVLDKIKLNKONILIRDKRFHY---DRNNIAGADESVVKEAHREVINSSTGL-- 627
DB 599 KIYV-----KLRGMNLIKTPYFTNFDDYNNYP--STWSNVNTNQDGLQGSANKLNG 652
QY 628 -----LLNIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656
DB 653 ETKIKIPMSSELKPKRYVFGYSKDPULTSIIIVKAKKEKTDYLVPEQGYTKFSYEFE 712
QY 657 NDRYDMLNIS-SLRQDCKTIDFKYND--KLPLYISNPYKV-----NVYAVTKENT 706
DB 713 TTEKDSNIEITLIGSGTYYLDNLISITELNSTPEILDEPEVKIPTDQEIIMDAHKIYFADL 772
QY 707 IINPSENGDTSTNGI 721
DB 773 NFNPS-TGNVINGM 786

RESULT 4
ID O06498 PRELIMINARY; PRT; 879 AA.
AC O06498;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Sb component.
OS Clostridium spiroforme.
OC Bacteria; Firmicutes; Mollicutes.
OX NCBI_TaxID=29348;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS246;
RA Gilbert M., Perelle S., Daube G., Popoff M.R.;
RT "Clostridium spiroforme toxin genes are related to C. parvifragens iota
RT toxin genes but have a different genomic localization."
RL Syst. Appl. Microbiol. 20:337-347(1997).
DR EMBL; X97569; CA666612.1; -.
DR HSSP; P13423; 1ACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxin; I.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 879 AA; 98739 MW; 40685ACB8E05BA01 CRC64;
```

```
Query Match 23.3%; Score 877.5; DB 2; Length 879;
Best Local Similarity 28.7%; Pred. No. 2.8e-33;
Matches 259; Conservative 147; Mismatches 259; Indels 237; Gaps 34;

QY 4 QENRLNESSSOGLLYGYSFSLNFCAPMVVTSSTGDLSPSSELENIPSENQ-YFQS 62
DB 36 QEVTTREKTVSSNGLMGYFADHFKDLMLAPVKNKRGELKPEKNKVEKULTEKTNKS 95
QY 63 AINSGFTKVKSDYEYFATSADNHVTMWYDDQEVINKASNKIRLEKGRLYQIKIQOR 122
DB 96 IRWTGRIIPSKDGEYTLSTDKDN-VLMQINAE---GEIANTLVNMKGQEYSIRIEIQ- 150
QY 123 ENPTEKGLDF-----KLYWDSQNKKEVISDNLOLPELKQKSSN----- 162
DB 151 ----DKDIGYVDDLSSPKLYW-ELNGDKTLPIPEKNFLRDYSKIDENDPFPKDNFFDLK 205
QY 163 --SRKKESTSA-GPTVPDRDNDGIPDSLEVEGYTVDVKNKRTFLSPFWSNHEKKGILTKY 219
DB 206 LKSSARLASGWDGDELDTDNDNIPDAVEKNGYTI----KDSIAVKW-EDSFAQQGYKY 260
QY 220 KSSPEKWSADPSYDFEKTGRIDKNGSPEARHPLVAAYPIVHVDMENILSKNEDOST 279
DB 261 LSSYLESTAGDPTDYQKASGSFKAITEARDPLVAAYPIVGVGMKLLISTNEHAST 320
QY 280 QNTDSETRTTSKNTSTSRTHTEVHGNAEVHASFDDIGSVSAGFSNNSSTVAIDHS 337
```

```
Db 20 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSLENIENISENQYF 79
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 80 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 139
QY 121 QRENTEKGLDFKLYWTDSONKKEVTSNDLQPLKOKSSNSRKRSTSGPTVPDRDN 180
Db 140 QRENTEKGLDFKLYWTDSONKKEVTSNDLQPLKOKSSNSRKRSTSGPTVPDRDN 199
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240
Db 200 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 259
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 300
Db 260 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 319
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 320 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 379
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 380 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 439
QY 421 LNAQKFSSTPIITMNYNQFLEKTKQLRDLTDQVYGNATYNFENGVRVVDTSKNWSEV 480
Db 440 LNAQKFSSTPIITMNYNQFLEKTKQLRDLTDQVYGNATYNFENGVRVVDTSKNWSEV 499
QY 481 LPOIQTETARIIPNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATFGNPNGL 540
Db 500 LPOIQTETARIIPNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATFGNPNGL 559
QY 541 QYQKDIETEDFDFDQSTQNIKNQLAELNATNIYVLDKIKNKAKNMLIRDKRPHYDR 600
Db 560 QYQKDIETEDFDFDQSTQNIKNQLAELNATNIYVLDKIKNKAKNMLIRDKRPHYDR 619
QY 601 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 660
Db 620 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 679
QY 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 680 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 739
QY 721 IK 722
Db 740 IK 741
```

## RESULT 2

```
Q937W3
ID Q937W3 PRELIMINARY; PRT; 743 AA.
AC Q937W3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE pag protein (Fragment).
GN PAG.
OS Bacillus anthracis.
OG Plasmid pXO1.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Carbosap;
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,
RA Fasanella A., Francia M., Ciuchini F.;
RT "Sequence analysis of the genes encoding for the major virulence
RT factors of bacillus anthracis vaccine strain 'Carbosap'.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL: AJ413936; CAC93934.1; -.
DR GO: GO:000576; C:extracellular; IEA.
DR GO: GO:0046821; C:extrachromosomal DNA; IEA.
DR GO: GO:0015070; F:toxin activity; IEA.
DR GO: GO:0009405; P:pathogenesis; IEA.
DR InterPro: IPR003896; Anthrax toxinB.
DR Pfam: PF03495; Binary toxB; I.
DR PRINTS: PRO1391; BINARYTOXINB.
DR Plasmid.
KW NON_TER.
FT NON_TER 1 743
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5DE5A6 CRC64;
```

Query Match 97.4%; Score 3676; DB 2; Length 743;

Best Local Similarity 99.2%; Pred. No. 1e-164;

Matches 716; Conservative 1; Mismatches 0; Gaps 0;

```
QY 1 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSLENIENISENQYF 60
Db 22 EVKQENRLNSESSESSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSLENIENISENQYF 81
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 82 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 141
QY 121 QRENTEKGLDFKLYWTDSONKKEVTSNDLQPLKOKSSNSRKRSTSGPTVPDRDN 180
Db 142 QRENTEKGLDFKLYWTDSONKKEVTSNDLQPLKOKSSNSRKRSTSGPTVPDRDN 201
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 240
Db 202 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASTASDPYDFEKT 261
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 300
Db 262 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSTRISKNSTSRHT 321
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 322 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 381
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 382 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 441
QY 421 LNAQKFSSTPIITMNYNQFLEKTKQLRDLTDQVYGNATYNFENGVRVVDTSKNWSEV 480
Db 442 LNAQKFSSTPIITMNYNQFLEKTKQLRDLTDQVYGNATYNFENGVRVVDTSKNWSEV 501
QY 481 LPOIQTETARIIPNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATFGNPNGL 540
Db 502 LPOIQTETARIIPNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATFGNPNGL 561
QY 541 QYQKDIETEDFDFDQSTQNIKNQLAELNATNIYVLDKIKNKAKNMLIRDKRPHYDR 600
Db 562 QYQKDIETEDFDFDQSTQNIKNQLAELNATNIYVLDKIKNKAKNMLIRDKRPHYDR 621
QY 601 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 660
Db 622 NNIAGADESVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIETEDTGLKEVINDRY 681
QY 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 720
Db 682 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 741
QY 721 IK 722
Db 742 IK 743
```

## RESULT 3

O32739

ID O32739 PRELIMINARY; PRT; 876 AA.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds  
(without alignments)  
6697.424 Million cell updates/sec

Title: US-09-848-909A-11  
Perfect score: 3774  
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILFSSKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRMBL 25:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_rvirus:  
16: sp\_bacteriap:  
17: sp\_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3676	97.4	741	2 Q937W2	Q937W2 bacillus an
2	3676	97.4	743	2 Q937W3	Q937W3 bacillus an
3	877.5	23.3	876	2 Q32739	Q32739 clostridium
4	877.5	23.3	879	2 Q06498	Q06498 clostridium
5	875.5	23.2	876	2 Q9KH41	Q9KH41 clostridium
6	871	23.1	875	2 Q46221	Q46221 clostridium
7	813	21.5	721	2 Q86171	Q86171 clostridium
8	721	19.1	775	2 Q844J8	Q844J8 bacillus th
9	296.5	7.9	225	2 Q8KYK2	Q8KYK2 bacillus an
10	226.5	6.0	1881	16 Q8RGK2	Q8RGK2 fusobacteri
11	209.5	5.6	3468	5 Q81104	Q81104 plasmodium
12	202	5.4	4688	16 Q9PQ08	Q9PQ08 ureaplasma
13	201.5	5.3	2849	5 Q81HY4	Q81HY4 plasmodium
14	200.5	5.3	1904	5 Q81KS2	Q81KS2 plasmodium
15	196.5	5.2	2940	5 Q81HP9	Q81HP9 plasmodium
16	194	5.1	1404	2 Q8RJN9	Q8RJN9 mycoplasma

## ALIGNMENTS

## RESULT 1

Q937W2 PRELIMINARY; PRT; 741 AA.  
ID AC Q937W2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Pag protein (fragment).  
GN PAG.  
OS Bacillus anthracis.  
OG Plasmid pXO1.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Feetara;  
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M., Paganella A., Francia M., Ciuchini P.;  
RT "Sequence analysis of the genes encoding for the major virulence factors of Bacillus anthracis vaccine strain 'Carbosap'";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ413937; CAC93935.1;  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax\_coxinB.  
DR Pfam; PF03495; Binary\_coxB; I.  
DR PRINTS; PR01391; BINARYTOXINB.  
FT NON\_TER 1 741  
FT NON\_TER 741 741  
SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 97.4%; Score 3676; DB 2; Length 741;  
Best Local Similarity 99.2%; Pred.No. 1e-164;  
Matches 716; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESQGLLYFSDLNFLQAPMVVTSSTGDLSPSELENIPSENYQF 60

Q897H6 clostridium  
Q98PM9 mycoplasma  
Q81L45 plasmodium  
Q81548 plasmodium  
Q81FP4 plasmodium  
Q814R2 plasmodium  
Q8NWQ6 staphylococ  
Q8WSK5 plasmodium  
Q26216 plasmodium  
Q7YWE7 plasmodium  
Q7YWE6 plasmodium  
Q81G0 plasmodium  
Q81GX1 bacillus ce  
Q25579 helicobacte  
Q815A6 plasmodium  
Q9GZ76 plasmodium  
Q7YWF0 plasmodium  
Q81I47 plasmodium  
Q81JX9 plasmodium  
Q9AV25 oryza sativ  
Q7XC10 oryza sativ  
Q81C27 plasmodium  
Q813R5 plasmodium  
Q99U54 staphylococ  
Q931R6 staphylococ  
Q81H2 plasmodium  
Q81I18 plasmodium  
Q81B09 plasmodium  
Q81531 plasmodium

Db 537 LKGEIQTEISNBEHLSQSQTAAAEKAAVAATNNELSEK--NSLQTL-CNA----- 585

Qy 371 TAPIYVLPPTTSVLGKN-QTLATIDADENQJS---QILAPNNYPPSKNLAPIALNAOKK 426

Db 586 ----FQEKLAQSVQKLKENEQNFSSLDTSFKLNESHQBLENNHQITIKQL-----KD 634

Qy 427 FSTPTITMYNOFLELEKTKQLRLDQDVYGNIAYNFENGVRVDTGNNWSEVLPIQOE 486

Db 635 TSS-----KLQQLERANPEQKESTLSDENNDLRTK-----LLKLEE 672

Qy 487 TTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDW-TLKEALKI----AFGFNEPENGNIQ 541

Db 673 SNKSLIKKQEDVDSLEKNI-----QTLKEDLRKSEEARFSEAKNLREVIDNL- 722

Qy 542 YQCKDIT-EFDNPDQOTSQNTKNQLAELNATNIYTVLDKIKLNAMMILIRDRPHYDR 600

Db 723 -KGKHETLEAQRNDLHSSSLSDAKNTNAILSSBLTKSSESDVKELTANVETLTQDSK----- 776

Qy 601 NNIAVGADSVVKEAHREVINSGTEGLLNIDKDIR-----KILSGVIVEIE-----DT 649

Db 777 -----AMKQSTSLVNSYQS--ISNLYHELRDDHVNMQSQNNYLLSESKLTDC 824

Qy 650 EGLKE---VINDRYDMLNSSLRQDGKTFIDPKKYNKDLPLYISNPNYKVNYYAVTKENT 706

Db 825 ENLTQQNMTLIDNVQKLMHKHVNQESKV-SELKEVNGKLSLDLKNLRSLLAV-AISDNDQ 882

Qy 707 IINP-----SENGDT-----STNGIKKI 724

Db 883 ILTQLAELSKNYDSLEQESAQLNSGLKSL 911

Search completed: May 3, 2004, 19:36:59  
Job time : 12.2539 secs



FT CONFLICT 887 887 T -> P (IN REF. 2).  
FT CONFLICT 1284 1284 Y -> S (IN REF. 2).  
SQ SEQUENCE 1420 AA; 160000 MW; 7F6CF4BBE0FAC918 CRC64;  
  
Query Match  
Best Local Similarity 4.2%; Score 159; DB 1; Length 1420;  
Matches 155; Conservative 129; Mismatches 287; Indels 242; Gaps 40;  
  
QY 27 LNFQAPMV-VTSTGTCGLSIPSELEINSENOYFQSAL--WSGFIKVKSDY----- 77  
DB 274 LIFAQPVADIENTSGD-----PSEFHCLODALDAIDDFIQLKTAAYRTPOSS 322  
QY 78 -----TFATSDAHNVHTVMVDDQEVINKASNKIRLEKGRLYQI-----KIYQREN 124  
DB 323 GVLSSNIAGTNPLSSDGYT-----EQFQHYKN-NSISSQSPASVHSQETNKISPKDPS 375  
QY 125 PTEKGLDFKLYTDSQNKKEVVISDMQLQPLKQKSSNRKRSSTAGTVPDRNDGIP 184  
DB 376 PNF7GTGDKMLSPSDQAFAPLTPNNNINE--NELEFNDRKQTIVS-----NDLENSPLK 428  
QY 185 DSLEVEGYTVVKNKRTFTSPWISNIHEKKGTIKVKSPEKW-STASDPYSDFE----- 237  
DB 429 TELEANGRSLEKVNNSVKTGSDVTLNKEGTLEQEQENELPSDKSDSWDKELFGDE 488  
QY 238 -----KVTGRIDKNVGEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSET 286  
DB 489 DEDLFGDSNKSNTSNEKSISDEI-----TDMFEMSDEEENNN----- 529  
QY 287 RTISKNTSTSRTHTSVHGNAEVHASFDDIGGSVAGRSNSSTVAIDHSLSLAGERTW 346  
DB 530 KSIKN-----NKENHT--DLGKDIPF-FPSSSEKPI-----RTM 561  
QY 347 AETMGLNTADTAELANIRYVNTGTAPIYV-LPTSLVLGKKNQTLATIDADENQSOIL 405  
DB 562 SGT-----TKLNGKRKYLD--IPIDEMTLPTSPLYMDPGAPL-PVETPRDRKRSVF 610  
QY 406 APNNYYP-SKNLAPIALNAQKFSSTPI-----TMVYN-QFLEKTKQLRLTDQVYGN 459  
DB 611 APLNFPILNNDVKNYKSGGKFSFSPLOKEERALNFDISWADLSSEE--BEDDEENG 667  
QY 460 ATYNFENGRV-----VDTGSNWSE-----VLPQIOETTARII-----FNG-KDIN 499  
DB 668 SDEDLKSLNVRDMDKPSDNISTNTNIHEPOYINYSIPSLQDS---IIKOENFNSVNDAN 724  
QY 500 LVPER-----IAAVNSDPLETTKPDWTLKEALKIAPFNENGNLQYQKGIITFEDNF 554  
DB 725 ITSKEGFSIWKIPQNDIPQTESPLKTVDSII----- 757  
QY 555 DQTSQNIKNQLAELNATNIYTVLDIKLNAKNILIRDKRFHYDRNNIAGDSVWKE 614  
DB 758 QPIESNIKMTLEDNNVTSNPSEFTPMVNSBISNLPKKS-----GIPEFTPAD 806  
QY 615 AHREVINSTEGLLL-----NIDKQIRKILSGYIV---EIEDT 649  
DB 807 PNLSPFESSSLPFLRHMP-LASIPDIFITPTPVVTISEKEQDILDIAEQVVTYDYNILGN 866  
QY 650 EGLKEV-----INDRYDMLNLTSSLRQDKTFIDFKYNDKLPYISNPYKNVYAVTKEN 705  
DB 867 LGIPKIAVGVKCOQEGLIITWQLFSTP-DRINGNDITISKEY---NMKQPVVFKKH 922  
QY 706 TIINPSNGD-----TSTNGIK--KILIFS 728  
DB 923 ELIKVKHDSQFFIKFLNFRPPGINKPKSLLS 955  
  
RESULT 14  
ID CAGA\_HELPJ STANDARD; PRT; 1167 AA.  
AC Q9ZLTI;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cytotoxicity associated immunodominant antigen (120 kDa protein)

DE (CAGA pathogenicity island protein 26).  
GN CAGA OR CAI OR CAG26 OR JHP0495.  
OS Helicobacter pylori J99 (Campylobacter pylori J99).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteraceae; Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923682;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
RA Smith D.R., Noonan B., Guild B.C., deJonge B.J., Carmel G.,  
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.;  
RT "Genomic sequence comparison of two unrelated isolates of the human  
gastric pathogen Helicobacter pylori";  
RL Nature 357:176-180(1999).  
CC -!- FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,  
CC OR FUNCTION OF THE CYTOTOXIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC  
CC EMBL; AB001483; AB006073.1; .  
DR PIR: B71924; B71924.  
DR InterPro; IPR005169; Caga.  
DR InterPro; IPR004355; IVSec\_caga.  
DR Pfam; PF03507; Caga; 1.  
DR PRINTS; PRO1553; TYPE4SSCAGA.  
KW Antigen; Complete proteome.  
FT DOMAIN 246 249 POLY-THR.  
FT DOMAIN 882 889 POLY-ASN.  
SQ SEQUENCE 1167 AA; 129729 MW; FDE5E86B81CEBD02 CRC64;  
  
Query Match 4.2%; Score 158.5; DB 1; Length 1167;  
Best Local Similarity 19.6%; Pred. No. 1.3;  
Matches 157; Conservative 145; Mismatches 313; Indels 187; Gaps 41;  
  
QY 3 KOENRLNSESSESSQGLL---GYVFDL-NF-----QAPMVVTSSTGSLSPSSSE- 51  
DB 480 KADKALDREKTTIQLGSLKHGVMFVDYSNFKYTNASKSPDKGVGATNG-----VSHLEA 534  
QY 52 -----NIPSENQYFQSALWSGFIKVKSDYTFATSDAHNVHTVMVDD-----QEVIN 98  
DB 535 GFSKVAVENFLNMLNLAITSV---VRQLEDKLIAGKLSPOEANKLVKDFLSNKLVG 590  
QY 99 KASNSKIRLE---KGRLYQIKIYQRENPTKGLDFKLYWTDSONKKEVISSDNLQPE 155  
DB 591 KALNFKAVAEAKNTGYDEVK---QAQKLEKSL-----KKRELEKDVAK--N 635  
QY 156 LKQKSSNRKRSSTAGTVPDR-----DNDGIPDSLEVEGYTVVKNKRTFTSPWISNI 210  
DB 636 LESKSGNKNKMEAKSQANSQKDEIFALINKAENRDAEI-AYAQNLRKGIKRELSDKLENI 694  
QY 211 HEKGLTKYKSSPEKWSSTASDPYSDFE-----VTGRI-DKNVSPPEARHPLVAAYPIV 262  
DB 695 N--KDLKDFKSFDEFKNGKN--KDFSAEETLKALGSKVKDLGINFENIS----- 741  
QY 263 HVDMENILSKNEDQSTQNTDSETRTISKNTSTSRHTSEVHGNAEVHASFDDIGGSVSA 322  
DB 742 --KVENLNAALNEFKNGKNKDFSKVTQAK--SDLENSIKDVIINOKITDKVDNLNQAQSV 797  
QY 323 GFSNSNSSTVAIDHSLSLAGERTWAEITMGLNTADTAELANIRYVNTG-TAPIYVNLPTT 381  
DB 798 AKATGDFSGV-----EQALADLNKFSKEQLAQQAQKNEFDFTKKSALYQ----- 842  
QY 382 SILVKGKNTLATIDADENQLSQILAPNNYPSKNLAPIALNAQKFSSTPTTMYNNOFLE 441



QY	104	NKIRLEK	-----GRLYQIKIQORENPTEKGLDFKLVWTDSONKKEV	145
Db	365	SQTNVQRNTHYSRHYLPVPFANILDDNIYTIKRGPF	-----NLTNKGFNIE	409
QY	146	ISSDNLQL-PELKQKSSNRKKRSAGTVPDRDNDGIPDSLEVEGYTVVQNKKTFLS	204	
Db	410	NSGQNIERNPALQKLSSESWDLFTKVCURLTKGNSRD	-----DSTCIKVKQNL---	458
QY	205	PAISN-----THEKKGLTKYKSSPEKSWTASDPYS-DPEKVTGRIKDNVSPEARHPLV	256	
Db	459	PYVADKDSIQEIFENKIITD-ETNVQNY-----SDKESLDESILDGQVP--INEIIVDPL-	511	
QY	257	AAYPIVHVDMENIILSKNE-----DOSTQNTD-----SETRTISKNTS--TERTHTSEV	303	
Db	512	-----LNNVNEPLNLPGEEIVFYDDITKVDVLYNSYYVLESQKLSNNVENITLITSVEEA	567	
QY	304	HGNAEVHASFF-----DIGGSVSAFSSNSSSVAJDHSLSLAGERTWASTM-----	350	
Db	568	LGSNKIYTFLP9LAEKVNGKVGQAGLFLWANEVVEDFTNINKKOTLDKISDVSVIIPY	621	
QY	351	---GLNLTADTA-RLNANIRYVNTG-----TAPIYVNLPTTSLVIGKNQTLATID	395	
Db	628	IGPALNIGNSALRGNFQAFATAGVAFLLGEGFPEFTIPALGVFTFYSSIOEREKIITIE	687	
QY	396	-----ADENOLSOILAPNNY-----YPSKNLAPIALNAO-----KKFSS	429	
Db	688	NCLQQRVKRWKDSYQMWVSNLSRITTFQPHINYOYMSLSYQADAIKAKIDLEYKYSY	747	
QY	430	TPITMNYONFLEBKTKQRLRDLTDQVYGNF-----ATYFNENGRVRVDTGSNWSEVL	481	
Db	748	SDENIKSQVENIKNSLDVKI--SEAMNNINKFIRECSVTYLFKN-----ML	792	
QY	482	POIQETTARIIFNGKD--LNL-----VERRIAAVNPSPDLETTKPDMTLKEALK	528	
Db	793	PKVIDELNKFDLTKTELINLIDSHNLIILVGEVDRUKAKVNES--PENTMP-----	841	
QY	529	IAFG-FNEPENGNILOYGKQDITFDFPFDQOTS-----QNIKNQIAELNATNI-YTVLDKIK	582	
Db	842	--ENIFSNTNSLL--KDIINEYFNSINDSKILSQNKKNALVDTSYGNAEVVRGDNVQ	896	
QY	583	LNAKMWILIRDKPHYDRNNIAVCADESVVKEAAREVINSSTEGLLINIDKIRKILSGY	642	
Db	897	LN---ITYTNDFKLSSGDKIIVLNNNNILYSA--IYENSSVSFWIKISKDLTNSHNEY	950	
QY	643	IV-ETIEDTGLKEVINDRYDMNINISLRQDGKTFDFKYNKDKLPIY----ISNPNTKVN	697	
Db	951	TIINSIEQNSGWKLCLANG---NIEWILOPVN-----RKYSLIFDYSLSHTGYTNK	1001	
QY	698	VYAVTKENTII 708		
Db	1002	WFFWTITNNIM 1012		

RESULT 12

YJL3 YEAST

ID	YJL3 YEAST	STANDARD;	PRT;	1803 AA.
AC	P47024; P87192;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Transposon Ty4 207.7 kda hypothetical protein.			
GN	TY4B OR YJL113W OR J0750.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxId=4932;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=S288C / F1679;			
EX	MEDLINE=97103775; PubMed=948101;			
RA	Czepluch C., Kordes E., Fujol A., Jauniaux J.-C.;			
RT	"Sequencing analysis of a 40.2 kb fragment of yeast chromosome X			



RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Welfjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Moore K., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
RT "The genome sequence of *Schizosaccharomyces pombe*,"  
RL Nature 415:871-880 (2002).  
RN [2]  
RP IDENTIFICATION, AND GENE NAME.  
RX MEDLINE=20143585; PubMed=10679021;  
RA Jin H., Amberg D.C.,  
RT "The secretory pathway mediates localization of the cell polarity  
RT regulator Alp3p/Bud6p,"  
RL Mol. Biol. Cell 11:647-661 (2000).  
CC -!- FUNCTION: Involved in the organization and/or function of the  
CC actin cytoskeleton.  
CC -!- SIMILARITY: TO YEAST BUD6.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC ENBL; Z97208; CAB10112.1; --  
CC ENBL; AL109770; CAB52420.1; --  
CC GeneDB\_Spomb; SPAC15A10.16; --  
CC GO; GO:0005884; C:actin filament; ISS.  
CC GO; GO:0016043; P:cell organization and biogenesis; ISS.  
CC InterPro; IPR005613; AIP3.  
CC Pfam; PF03915; AIP3; 1.  
CC Coiled coil, Cytoskeleton.  
KW Coiled coil, Cytoskeleton.  
FT DOMAIN 1009 1096  
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40CC4FF7537 CRC64;  
Query Match 4.3%; Score 163; DB 1; Length 1385;  
Best Local Similarity 19.1%; Pred. No. 0.96;  
Matches 146; Conservative 129; Mismatches 292; Indels 198; Gaps 33;  
QY 26 DLNFQAPMVVTSSTGDL-----SIPSELENIPSENYQFQSAIWS-----GFIKVKKS 74  
DB 326 DRKFSPKLRTPSLTKSLDPTGPTSLKPSLRKSPS--SFVQKVYRSNLSLRISQANRS 383  
QY 75 DEVTATSDADNHVTMVVDQEVINKAS-----NSN 104  
DB 384 NVFPGATD---VTRSVSHRILSSSTINDGEVAPLPORSTISPSNPISATVLPST 439  
QY 105 KIRLEKGRIVQIKIYQREN-----PTEKGLDFKL-----YWTDSQNK 143  
DB 440 PILLPRGRSSTLSVKNKQFNADGSLNSPNSIRETEEYVAGPKLEADIADVEVTDATSR 499  
QY 144 EVI-----SSDNLQLPELKQSS-----NSRKKESTAGTVPDRNDGI-----PDSL 187  
DB 500 ELLERIOQAESSEDTSEISLQGGKUSLPQVSTQOETQPSSVPEAASNEIAKEPAVT 559  
QY 188 EYEGYTVDVKNKRTFLPSWISNHEKKGTIKYKSSPEKMASTADSDYDFEKTGTGRDKNV 247  
DB 560 AIESIT-----ERKEAPVISSEKIESGTS-STSDTKGGLANFENSLSEELERLIQNN- 613  
QY 248 SPARPLVAAPYIVHVDNENILSKNEDQSTQNTDSERTISKNTSRTSRTHSEVHNA 307  
DB 614 NAEQDEPSYFKHYKESSE-----SGSEDEFKSEKDTKGYVISNDSDTQVEDESDKSTP 669

QY 308 EVHASFDFDIGSVSAGFSNSNSTVAID-----HSLSLAGERTWAETMGLNTADTARLNAN 363  
DB 670 NTGA-----SAKLNDPSSTITVSDVYKPKASPVEITEPPSSALVSATSPTTNPV 720  
QY 364 I--RYNTGTAPYVNVLPSTSLVLGKNOTLATDADENQLSQI--LAPNNYYSKNLAPI 419  
DB 721 IVPFAVHLSA--FSTAPVSTIV-----SNISPLPTVAPPNVSGSPSETPI 764  
QY 420 ALNAQKFSSTPTMTNQNQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSE 479  
DB 765 S-----KPEKVPVVSQTEKAL-----PKPLGVDTKEYY--FLRYNNQTRKVKVSPSLNAN 813  
QY 480 VLPOIQETIARIIPNGK--DLNLVERIAAVNSDPLEITKPDMT-LKEALKIATFGENEP 536  
DB 814 ELGELFSNVYKISFGSDSYELNIED-----PDTKISYLLLEDLSLKYKSLVSPMFKEQ 866  
QY 537 NGNLQ--YQKDIETDFDNFDOOTSONIKNQLAELNATNIYT-VLDKIKLNAMNILIR 592  
DB 867 DANKKRDPHSGEVSALQHSQAQNTLDDHVNTTTHSPSSAFTTEILERLKA----- 917  
QY 593 DKRPHYDRNNIAGDESVVYKEAHRVINSSTGLLNLINIDKDKIRKILSGVIV-EIETEG 651  
DB 918 -----IEQNIST-----NHTNDSAAALKSSEDSHKLANNFSPVPSIDHKF 956  
QY 652 LKEVINDRYDMLNITSSLRQDGKTFI-----DFKK-----YNDK 684  
DB 957 YQVQVNMQ-----LELASLKQISAFAFTTRIPLKIKDFKKEINAFNEK 998  
RESULT 11  
BXD CLOBO  
ID - BXD CLOBO STANDARD; PRT; 1276 AA.  
AC P19321;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)  
DE (Bontoxilysin D).  
GN BOTD.  
OS Clostridium botulinum.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
CX NCBI\_TaxID=1491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BVD/-3;  
RX MEDLINE=91016853; PubMed=2216736;  
RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,  
RA Kozaki S., Kriegstein K., Henschen A., Gill D.M., Niemann H.;  
RT "Nucleotide sequence of the gene encoding Clostridium botulinum  
RT neurotoxin type D,"  
RL Nucleic Acids Res. 18:5556-5556 (1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CB16;  
RX MEDLINE=93042276; PubMed=1420572;  
RA Sunagawa H., Ohyama T., Watanabe T., Inoue K.;  
RT "The complete amino acid sequence of the Clostridium botulinum type D  
RT neurotoxin, deduced by nucleotide sequence analysis of the encoding  
RT phase d-16 phi genome,"  
RL J. Vet. Med. Sci. 54:905-913 (1992).  
RN [3]  
RP PARTIAL SEQUENCE.  
RC STRAIN=D-SA, and D-1873;  
RX MEDLINE=89339741; PubMed=2668193;  
RA Morishiki K., Syuto B., Kubo S., Oguma K.;  
RT "Molecular diversity of neurotoxins from Clostridium botulinum type D  
RT strains,"  
RL Infect. Immun. 57:2886-2891 (1989).  
RN [4]  
RP IDENTIFICATION OF SUBSTRATE.  
RX MEDLINE=94230352; PubMed=8175689;  
RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,

```
DR Subtilist; BG10797; wapa.
DR InterPro; IPR003305; CBM_Cenc.
DR InterPro; IPR006530; YD_Cenc.
DR Pfam; PF02018; CBM_4_9; 1.
DR Pfam; PF05593; Rhs_repeat; 14.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 17.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 17.
DR Cell wall; Repeat; Signal; Complete proteome.
FT SIGNAL 1 28 WALL-ASSOCIATED PROTEIN.
FT CHAIN 29 2334 3 X 101 AA APPROXIMATE TANDEM REPEATS.
FT DOMAIN 504 869
FT REPEAT 504 605 1-1.
FT REPEAT 636 736 1-2.
FT REPEAT 769 869 1-3.
FT DOMAIN 1021 2139
FT REPEAT 1021 1040
FT REPEAT 1042 1061
FT REPEAT 1063 1082
FT REPEAT 1083 1102
FT REPEAT 1109 1128
FT REPEAT 1129 1148
FT REPEAT 1150 1169
FT REPEAT 1174 1193
FT REPEAT 1199 1218
FT REPEAT 1219 1238
FT REPEAT 1246 1265
FT REPEAT 1267 1286
FT REPEAT 1290 1309
FT REPEAT 1312 1331
FT REPEAT 1332 1351
FT REPEAT 1352 1371
FT REPEAT 1372 1391
FT REPEAT 1392 1411
FT REPEAT 1412 1431
FT REPEAT 1432 1451
FT REPEAT 1452 1471
FT REPEAT 1472 1491
FT REPEAT 1492 1511
FT REPEAT 1512 1531
FT REPEAT 1532 1551
FT REPEAT 1552 1571
FT REPEAT 1572 1591
FT REPEAT 1592 1611
FT REPEAT 1612 1631
FT REPEAT 1632 1651
FT REPEAT 1652 1671
FT REPEAT 1672 1691
FT REPEAT 1692 1711
FT REPEAT 1712 1731
FT REPEAT 1732 1751
FT REPEAT 1753 1772
FT REPEAT 1773 1791
FT REPEAT 1792 1811
FT REPEAT 1812 1831
FT REPEAT 1832 1851
FT REPEAT 1852 1871
FT REPEAT 1872 1891
FT REPEAT 1892 1911
FT REPEAT 1912 1931
FT REPEAT 1932 1951
FT REPEAT 1952 1971
FT REPEAT 1972 1991
FT REPEAT 1992 2011
FT REPEAT 2012 2031
FT REPEAT 2032 2051
FT REPEAT 2052 2071
FT REPEAT 2072 2091
FT REPEAT 2093 2112
FT REPEAT 2120 2139
SQ SEQUENCE 2334 AA; 258329 MW; B75138CCD278BA3 CRC64;

Query Match
Best local similarity 18.2%; Score 164.5; DB 1; Length 2334;
Matches 173; Conservative 126; Mismatches 336; Indels 317; Gaps 42;

QY 1 EVKQENLLSESSOGILGYFSDLNFOA-----PMVVTSTTGDLSPSSSELENIP 54
D 269 EVERSDFKSVKLENEEGYLLHLTADENWLKDPERYVPISIDPSTLSVSSDTFVMSAYP 328
QY 55 SENQYFOAIWSGFIK--VKKSDEYTPATSADNHVTM----- 89
D 329 TTNYSASSQWDMANLKAAYLKTGYD-KTGTNYAFMKFNKLPIONMTVTKATLKTYVA 387
QY 90 -----WDDQEVINK-----ASNSKIRLEKGR--LYQIKIQOR 122
D 388 HSYGTGKATGLWLD---TVNSVDNAKVTWNTKPAKNIQKADVHKQWASVDVTAUKS 444
QY 123 ENPTKGLDFKLYWTDSONK---KEVSSDNLQLPKOKSSNRKRKSTSGAGTVP--- 176
D 445 WNSGGANGYFKLH-TNGNGKEYWKKLIS-----SANSANKPYIEVYTIPTKGN 491
QY 177 -----DRNDGIPD-----SLB-VGQYVDVKNKETFPLSPWISNHEKKGI-TKYKSSP 223
D 492 TPTTKAYHNGSTGYFDLSWKKVEGAGYKWKYIYNGKEYQAIAGNV-----TSWSTKG 545
QY 224 EK-WSTASDPDYSDEKVTGTGRID-KNVSPPEARHPLVAAYPIVHVDMENIILSKNE----- 275
```

```
Db 546 KKIWPTSAEIAKRYKL--HLGKDAELALDP-----SPVYKNSGGYSATSKNYWIGVSA 599
QY 276 --DOSTQNTDSETRTISKNTSTSRHTSEVHGNAEVAHSPFDIGGVSAG-----PS 325
Db 600 IFDQEGGAMSAPAKPVPINVGKAQAFSAKYNNGNA--TCYFDLSWKAWSGATGYKQOVEN 658
QY 326 NSNSSTVAIDH--SLSLAGERTWAETMGLNTADTA-----RLNANIRYVNTGPA 372
Db 659 GKGFTLDLGNQTSWTTGKKIWIPTSAEIKAGKAYALHLKDGSGAELPINFPGTYKNAGGD 718
QY 373 PIYVNIPTSLVLGKQNTLATIDADENQLSOILAPNN-----YYPK-----NLA--- 417
Db 719 GAKENYSFKIIAYNKDGEAIAAPATPALPDIARPKNVGTGLYNTKSSQGTGVNLINEK 778
QY 418 -----PIALNAQKFFSTPI-----TMNYNOFLELEKX-----QLRL 450
Db 779 VQNAKGYKVIYNGKEYQSFVDCGADHWTTQNKNIWPTSEIKAGSVKYLHTDGKGELAL 838
QY 451 DTDOVTGN-----IATYN--PENGVRVDVTSNWSSEV-L 481
Db 839 DPGFVYNNANGYKGNYSFTLVAYDANGETPTAPFNPFFHEGAELFJGTEEYWSIIDI 898
QY 482 P--QIQTETARIIFNGKDLNL-----VERRIAAVNPSPDL-----ETT----- 517
Db 899 PSGOLNGATGVIVNBEEDLSIDGPGGLGSRYSNLSDDHLFGQGWYADAETSIVSTD 958
QY 518 -----KPDMTLKEALKIARG-FNEPENGNYQYQKGIITFDNFDOQTSONIKNQLAEL- 569
Db 959 GGAMYIDEDATTHRTFKKADGTYPPTGVYLETETADQFILKTKDQTNAYFNKGGKQLQ 1018
QY 570 -----NATNIYTVLDKIKLNAMNLIIRDKRPHYDRN-----NIAVGADSV 611
Db 1019 KVDGHNNAIVYTYNDKQNLTAITDASGRKLITTYDENGHVTSITGPKNKVTVSYENDL 1078
QY 612 VKE-----AHR--EVINSSTGLLL 629
Db 1079 LKVTDTDTGTVSYDYDSEGRLVKQVSAKVPFTEYQYSGHRLKAINAKKETYVY 1138
QY 630 NIDKDI-----RKILSGY-----IVEIETDEGLKEVINDRYDMLNI 665
Db 1139 SYDAKDKTLLMTQPNGRKQVGYNEAGNPQIVDDAEGLKITNTKIEGNV 1190

RESULT 10
FAT1_SCHPO
ID FAT1_SCHPO STANDARD; PRT; 1385 AA.
AC O13735; QOUT0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Swilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Feat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
```



```

Db 777 YQSVKFKYELRKNLGYMGHIVLIAEVVFKLYKVDYISRYSNLQTEWQYYSZVL 836
QY 700 AVTK-----ENTNIPSENGDTSN 719
Db 837 NETRMYSKILGGSYIDGNGNIIPQLPDNTVLTLP--NGDASN 880

RESULT 8
SMC3_YEAST
ID SMC3_YEAST STANDARD; PRT; 1230 AA.
AC P47037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Structural maintenance of chromosome 3 (DA-box protein SMC3).
GN SMC3 OR YJL074C OR J1049.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]_SEQUENCE FROM N.A.
RP STRAIN=W303;
RX MEDLINE=97474309; PubMed=9335333;
RA Michaelis C., Ciolek R., Nasmyth K.;
RT "Cohesins: chromosomal proteins that prevent premature separation of
RL sister chromatids.";
RN Cell 91:35-45(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96208490; PubMed=8641289;
RA Galibert F., Alexandraki D., Baur A., Boles E., Chalwatzis N.,
RA Chuat J.-C., Coster F., Cziepluch C., De Haan M., Domdey H.,
RA Durand P., Entian K.-D., Gatius M., Goffeau A., Grivell L.A.,
RA Hennemann A., Herbert C.J., Heumann K., Hilger F., Hollenberg C.P.,
RA Huang M.-E., Jacq C., Jauniaux J.-C., Katsoulou C., Kirchhath L.,
RA Kleine K., Kordes E., Koetter P., Liebl S., Louis E.J., Manus V.,
RA Mewes H.-W., Miosga T., Obermaier B., Perea J., Pohl T.M.,
RA Portetelle D., Pujoil A., Fumelle B., Ramezani Rad M., Rasmussen S.W.,
RA Rose M., Rossau R., Schaaf-Gerstenschlaeger I., Smits P.H.M.,
RA Scarcez T., Soriano N., Tovani D., Tzermia M., Van Broekhoven A.,
RA Vandenbol M., Wedler D.H., Von Wettstein D., Wambutt R., Zagulski M.,
RA Zollner A., Karpfinger-Hartl L.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RN X.";
RL EMBO J. 15:2031-2049(1996).
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1678;
RA Sor F.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[4]
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC1, IRR1 AND MCD1, AND
RP INTERACTION OF THE COHESIN COMPLEX WITH SCC2.
RX MEDLINE=99145468; PubMed=990856;
RA Toth A., Ciolek R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;
RT "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to
RT establish cohesion between sister chromatids during DNA replication.";
RL Genes Dev. 13:320-333(1999).
[5]
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3, MCD1 AND IRR1, AND
RP STRUCTURE.
RX MEDLINE=21980168; PubMed=11983169;
RA Haering C.H., Loewe J., Hochwagen A., Nasmyth K.;
RT "Molecular architecture of SMC proteins and the yeast cohesin
RT complex.";
RL Mol. Cell 9:773-788(2002).
CC -!- FUNCTION: Involved in chromosome cohesion during cell cycle and in
CC DNA repair. Central component of cohesin complex. The cohesin
CC complex is required for the cohesion of sister chromatids after
CC DNA replication. The cohesin complex apparently forms a large
CC proteinaceous ring within which sister chromatids can be trapped.

```

```

CC At anaphase, the complex is cleaved and dissociates from
CC chromatin, allowing sister chromatids to segregate.
CC SUBUNIT: Cohesin complexes are composed of the SMC1 and SMC3
CC heterodimer attached via their hinge domain. MCD1/SCC1 which link
CC them, and IRR1/SCC3, which interacts with MCD1. The cohesin
CC complex also interacts with SCC2, which is required for its
CC association with chromosomes.
CC -!- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin.
CC Before prophase it is scattered along chromosome arms. At
CC anaphase, the MCD1 subunit of the cohesin complex is cleaved,
CC leading to the dissociation of the complex from chromosomes,
CC allowing chromosome separation.
CC -!- DOMAIN: The flexible hinge domain, which separates the large
CC intramolecular coiled coil regions, allows the heterotypic
CC interaction with the corresponding domain of SMC1, forming a V-
CC shaped heterodimer. The two heads of the heterodimer are then
CC connected by different ends of the cleavable MCD1 protein, forming
CC a ring structure (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. SMC3 subfamily.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; Y14278; CAA74655.1; -
CC EMBL; Z49349; CAA89366.1; -
CC EMBL; X88851; CAA61313.1; -
CC PIR; S56850; S56850.
CC GERMONLINE; 141688; -
CC SGD; S0003610; SMC3.
CC GO; GO:0007151; P:sporulation (sensu Saccharomycetes); IMP.
CC GO; GO:0007130; P:synaptonemal complex formation; IMP.
CC InterPro; IPR003439; ABC transporter.
CC InterPro; IPR003405; SMC_C.
CC InterPro; IPR003393; SMC_N.
CC Pfam; PF02483; SMC_C; 1.
CC Pfam; PF02483; SMC_N; 1.
CC Mitosis; Cell cycle; Chromosome partition; ATP-binding; Coiled coil;
KW Nuclear protein.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT DOMAIN 172 482 COILED COIL (POTENTIAL).
FT DOMAIN 483 684 FLEXIBLE HINGE.
FT DOMAIN 685 1041 COILED COIL (POTENTIAL).
FT DOMAIN 1126 1161 ALA/ASP-RICH (DA-BOX)
SQ SEQUENCE 1230 AA; 141336 MW; B152D88F7780341F CRC64;

Query Match 4.5%; Score 170.5; DB 1; Length 1230;
Best Local Similarity 19.8%; Pred. No. 0.34;
Matches 182; Conservative 136; Mismatches 355; Indels 247; Gaps 38;

QY 1 EVKQENRLNESSSQGLLOYFSDLNFOAPMVVT--SSTGDLSTPSELENIPSENQ 58
Db 209 EMEQERLEKYNELERNKIYQFTLYDRELNEVINQMERLGDYN-----NTVVSSEQ 262
QY 59 YFQ-----SAIWSGFIKVKSDSEYTFATSNHVTWVDDQEVINK--- 99
Db 263 YIQELDKREDMDQVSKLSSI-EASLKIKNATDLOQAKLRSEISQKLNVANVKKIDVQ 321
QY 100 ---ASNKNIRLEKGRVYQIK-IQYQRENTEKGLDFKLYWTSQNKKEVSSDNLQPE 155
Db 322 QQIESNEEQRLNDSATLKEIKSIEQRKQKLSILPRYQELT---KEEAMY--KLQGLAS 375
QY 156 LKQKSSNSRKKRSTSGAGTVPDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIHE-KK 214
Db 376 LQQQRDLILKGYARFKSKDRDTHSEIE-----ELKSS-----IQNLNELES 422
QY 215 GLTKYKSPKSTASDPYSDFEKVTGRIDKNVSPPEARHPLVAAYPIVHVDMEIILSKN 274
Db 423 QLQMDRTSLRKQYSAID-----EIEELIDISINGPTKGQL-----EDFDSLEHLKQK 471

```

Db 810 STDKSTEDNTD---EKYFSALNTN---VTGDSCEIETASNVEE---NLRYCEKD 858  
Qy 547 ITEFENP-DQOTSON---IKNOLA-----ELNATHIYTVLDIKLNAMNLLIRDKR 595  
Db 859 MNEAEMSSGDECKQNDGSKTKQISFTSDPNFOESNDTEFSSTK-----905  
Qy 596 PHYDRNNIAGADESVVKEAHR-EVIN-----SSTEGLLINIDKD 634  
Db 906 ---YKVRNSLDEDESLLKELTAEVVDKLBDEESEDSEYQDYADPEPQNDGSGNENIVKG 963  
Qy 635 IKKILSGYIVETDEGLKEVINDRYDMLNIISSLRQDGKTFDFPKYNDKPLVLSNPNY 694  
Db 964 TKK-----DTLGIYEPENKVN-----KVHEEETLFEANVSS 995  
Qy 695 KVNVTAVTKENTIIN 709  
Db 996 SVNVQNKDHTDVIN 1010

## RESULT 7

S155 YEAST  
AC P436T2;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE S1T4-associated protein SAPI55.  
GN SAPI55 OR YFR040W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RN  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=96220458; PubMed=8649382;  
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,  
RA Arndt K.T.;  
RT "The SAP, a new family of proteins, associate and function positively  
RT with the S1T4 phosphatase."  
RL Mol. Cell. Biol. 16:2744-2755 (1996).  
[2]  
RN  
RP SEQUENCE OF 98-1000 FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=95400292; PubMed=7670463;  
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
RA Yamazaki M., Tashiro H., Eki T.;  
RT "Analysis of the nucleotide sequence of chromosome VI from  
RT Saccharomyces cerevisiae."  
RL Nat. Genet. 10:261-268 (1995).  
[3]  
RN  
RP SEQUENCE OF 98-1000 FROM N.A.  
RC STRAIN=S288C / AB972;  
RX MEDLINE=96287652; PubMed=8686379;  
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,  
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;  
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of  
RT chromosome VI from Saccharomyces cerevisiae."  
RL Yeast 12:149-167 (1996).  
CC  
CC -1- FUNCTION: Associates with the S1T4 phosphatase in a cell cycle  
CC dependent manner. May be directly or indirectly involved in  
CC S1T4-dependent functions in budding and in normal Gi cyclin  
CC expression.  
CC  
CC -1- PPM: Hyperphosphorylated in the absence of S1T4.  
CC  
CC -1- SIMILARITY: Belongs to the SAPS family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC  
DR EMBL; U50560; AAC49303.1; -;  
DR EMBL; D50617; BAA09279.1; -;  
DR Germline; 140194; -;  
DR SGD; S0001936; SAPI55.  
DR InterPro; IPR007587; SAPS.  
DR Pfam; PFO499; SAPS; 1.  
KW Phosphorylation; Cell cycle.  
FT CONFLICT 663 689  
FT CONFLICT 674 689  
FT  
FT  
FT CONFLICT 814 815  
FT CONFLICT 818 822  
FT  
SQ SEQUENCE 1000 AA; 114880 MW; AD4FBF0BC7C588D5 CRC64;  
Query Match 4.5%; Score 171; DB 1; Length 1000;  
Best Local Similarity 19.5%; Pred. No. 0.25;  
Matches 173; Conservative 130; Mismatches 321; Indels 262; Gaps 42;  
Qy 1 EVQENRLNSESSESSQGLLYYFS--DL-----NFQAPMVVTSSTGD-----LSIPS 47  
Db 90 EKEEERSMNSSESSTSFSSGSTSKTDLDEEDISNATAPMMVYTKNLNDSFIERLVLVET 149  
Qy 48 SELENTPEENQYFQSALWSGFIKVKKSDEYTFATSNADNHVWVDD-QEVINKASNSKI 106  
Db 150 ELNELSRQNKTLIDPFCGFFPKTKNKV-----NNWEYLVDQLMECISKIKTATTV 203  
Qy 107 RLEKGRLYQIKIQRENPTTEKGLDFKLYWTDSONKEVSSDNQLQPELKQKSSNRKK 166  
Db 204 DLNN-----LIDYQEQQLD-----DSSQEDVYVESDEQEKEEDDNNNNKK 247  
Qy 167 R-----STSAGPTVPDEDNDGIPDSLEVBGYTVDVKNKRTFLSP-----WISNIHEKKG 215  
Db 248 RRRKGSSEFGNDDINNDDDDANEDDESAYL-----TKATIISEIFSLDIWLIS-----ES 299  
Qy 216 LTKYKSSPEK-WSTASDPYSDFEKVTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKN 274  
Db 300 LVKNQSYLNKIWSIINQP-----NFNSE-NSPLV---DIFLQINQLLITR- 341  
Qy 275 EDQ-----STQNTDSET-----RTISKNT 293  
Db 342 QDQYLNFIPTERSFVDDMLKHVDISLLMDFLKIISTDKIESPTGIELVYDQNLISKCL 401  
Qy 294 STSRHTSEVHGNAEVAHFDDIGGSVAGFSNSNSTVAI-DHSLS--LAGERTWAETM 350  
Db 402 SFLNNKSP---ADIAQCVGDFLKALIAISANAPLDDISIGPNSLTRQLASPEIAKL 457  
Qy 351 GLNTADTARLNANIRYVNTGTAPIYVLPPTSLVLGKNQPLATIDADENQLSILAENNY 410  
Db 458 DI-----MINQGAALNT-----TVSIVI---ELIRKNSDYDQVNLITTTIKT 498  
Qy 411 YPSKNLAPIALNAQKFSSTPTMNNQFLEKTKOLRLDTQVYGNATYFNGRVR 470  
Db 499 HPPSNRDPYLYGLLRKFSNHLSDFFQIILDIENDANIPLHENQLHEKFKPLGFERKV- 557  
Qy 471 VDTGSNNSEVLPOQTETARIIFNGKDLNIVERRIAAVNP--SDPLETTKPMTL--KEA 526  
Db 558 -----VELIAELLHCNMGMLNSKRAERIARRRDKVRSQLSHLLQDALNDLSIEKEQ 610  
Qy 527 LKIAFG-----FNEPNGNLQYQKDI-----TEPDFNP-----QQTSONIKNOLAE 568  
Db 611 LKTKHSPTRTDHLKNNNGKINDNNDNDESDYGDIDEISFEIPIYNNKQNLKLTDP 670  
Qy 569 LNATNIYTVLDKIKLNARMNILIRDKRFHYDRNNIAVGAD-----ESVVKAEHREVIN-- 621  
Db 671 TVGT-----YKSNYMIKD-----CFQNNELFLTHPWNFWHNVIFDIQOIFNGR 716  
Qy 622 ---SSTEGLLINID--KDIKILSGYIVET---DTEGLKEVINDRYDMLNIISS--LR- 669  
Db 717 MDFSNSFLVLSLNLKSSYQFMTDIVISDEKGTDSRFSFVRDPNDFPKITTFILRG 776  
Qy 670 -QDGKGTIDFKKNY-----DKLFLYISN-----PNYKVNYY 699



DR EMBL: X02919; CAA26676.1; --  
DR PIR: A24594; A24594.  
DR PIR: S05603; S05603.  
DR PDB: 1CEJ; 28-MAY-99.  
DR InterPro: IPR006209; EGF\_like.  
DR Pfam: PF00008; EGF; 1.  
DR Malaria: Merzoite; Polyprotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor; 3D-structure.  
KW SIGNAL  
FT CHAIN 1 19  
FT CHAIN 20 1639  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 268 268 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 768 768 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 783 783 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 844 844 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 920 920 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1639 1639 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;  
Query Match 5.0%; Score 190.5; DB 1; Length 1639;  
Best Local Similarity 20.6%; Pred. No. 0.051;  
Matches 171; Conservative 138; Mismatches 255; Indels 265; Gaps 47;  
QY 5 ENRL-LNESSSSQGLGYFDLNFQAPWVTSSTTGLSPSELENIPENQVQSA 63  
DB 927 ENTLGKNNKYQELIGQSS-NF-----YKILKDSFTFYNE 965  
QY 64 IWSGFIKVKKSDEYTFATSDNHNVTWVDDQEVINKASN-----NKILEKGRLY- 114  
DB 966 SFINFVKSADD-----INSLNDESKRKLLEDINKLKTQLSPFLYNNKYKLERLFD 1020  
QY 115 -----QIKIQORENPTKGLDFKLYWTDQNKKEVIVSSDNLQLPELKQKSNRKRKS 168  
DB 1021 KKTGVGYKQIKKLLKEQLESK--NSLANPKHVL--QNFVFPFNKKEAEAEATEN 1076  
QY 169 TSAGTPVDRDNGIPDSLEVEGYTVDKNKRFLSPWISNIEKKGLTKY---KSSP-- 223  
DB 1077 T-----LENTKILLKH-----KGLVYKNGSSPLK 1103  
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAYPIHVVD 265  
DB 1104 TLSEESIQTEDNVASLENFKVLKLEGLKNDLNLEKKLSYLSGLHLIA-----E 1156  
QY 266 MENILSKNEDQSTQNTDSETRITKNTSTSRHTSEVHGNAEVHASFFDIGSVSAGFS 325  
DB 1157 LKEVIRKNN---YTGNSPSENNT-----DVNNALESYKKFLPEGTDTVATVVS 1200  
QY 326 NSNSSTVAIDHSLSLAGERTWAETMGLTADTARLNANIRYVNTGPTIYNVLPPT--- 381  
DB 1201 ESGSDTLEQSQKPKPASTHVGAES---NTITTSQ-NVDDVEDDVIIVPIFGSEEDYDL 1256  
QY 382 -SIVLGNQTLATIDADENQLSOLAPNNYPSKNLAPIA---LNQKXFSSTPTIMVYN 437  
DB 1257 GQVVTGEAVTSPVID---NILSKI---ENEYEVLYLKLPLAGVYRSLSKLENNVMTFVN 1310  
QY 438 -----QFLEKTKQLRLDTPQV-YGNIAT-----YNFNGRVRVDTGNSWSEVL 481  
DB 1311 VKDILNSRFKRNFKNV-LESDLIPYKDLTSSNVVVKPYKFLNKEKDKFLSNVYIK 1369  
QY 482 PQIQTETARIIFNGKDLNLVERRIAANVPSDPL-----ETTKPDM-TLKALKIAPGF 533  
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKSLDSIKKIYINDKQGE 1410  
QY 534 NEPN-----GNLYQOGKDIET-----FDNFDOQTSONIKNOLAELNATNIYT 576  
DB 1411 NEKYLPLANNIETLYKTVDNKIDLFVHLEAKVLNTYTK-----SNVEVKIKELN--YLKT 1465

QY 577 VLDKIKLNKAKON--ILIRDKRFHYDRNN-----IANGV-ADEVVVKEAHREVINSSTEGLL 628  
DB 1466 IQDKLADFKKNNNFVGIADLSTDYNNHNLTKFLSTGMVFENLAKTVLSNLDGNLQGG-M 1524  
QY 629 LNIDKD--IRKILSGYIVETEDTEGLKEVINDRYDMLNLISSLRQDQGTFFDFKKYNDKLP 686  
DB 1525 LNI5QHCVKK-----QCPQNSGCGFRHDE-----RECKCLLNYKQEGDKC- 1566  
QY 687 LYISNPYKNVYAVTENTIIINPENG-----DTSTNGIKKI 724  
DB 1567 --VENPNPTCN-----EN-----NGGCDADAKTTEEDSGSNG-KKI 1599  
RESULT 5  
P115\_MYCHR STANDARD; PRT; 979 AA.  
AC P41508;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE P115 protein.  
OS Mycoplasma hyorhinis.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2100;  
RN (1)  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=9113890; PubMed=1825306;  
RA Notariccola S.M., McIntosh M.A., Wise K.S.;  
RT "A Mycoplasma hyorhinis protein with sequence similarities to nucleotide-binding enzymes.";  
RL Gene 97:77-85(1991).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- DOMAIN: Consists of two putative central coiled-coil regions flanked by putative globular regions at the N- and C-termini.  
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYOPLASMA P115.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL: M34956; AAA25423.1; --  
DR PIR: JQ0894; JQ0894.  
DR InterPro: IPR003439; ABC transporter.  
DR InterPro: IPR005289; GTP-binding\_dom.  
DR InterPro: IPR003405; SMC\_C.  
DR InterPro: IPR003395; SMC\_N.  
DR Pfam: PF02483; SMC\_C; 1.  
DR TIGRFAMs: TIGR00650; MG442; 2.  
DR ATP-binding; Coiled coil.  
KW NP\_BIND 32 39 ATP (POTENTIAL).  
FT DOMAIN 169 224 COILED COIL (POTENTIAL).  
FT DOMAIN 231 400 COILED COIL (POTENTIAL).  
FT DOMAIN 569 821 COILED COIL (POTENTIAL).  
FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).  
SQ SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;  
Query Match 4.6%; Score 174.5; DB 1; Length 979;  
Best Local Similarity 19.1%; Pred. No. 0.16;  
Matches 177; Conservative 132; Mismatches 134; Indels 277; Gaps 40;  
QY 13 ESSQGLGYFDLNFQAPMVV-----TSSTGDLSPSELENIPSE 56  
DB 51 EQSAQLRGLNMDVIFAGSKTVKPKQKAMVLTFFKNEDAIEETKQFTTSRLKRGQGT 110  
QY 57 NOYFQS-----AIVSGFIKVKKSDEYTFATSDNHNVTWVDDQGE--VINKASN 102  
DB 111 NEYFNDQPVRYKDKNLAVESG---ISKSLAIISQGTISEIAETPEOKAVIEEAG 167

RT "Polymorphism of the precursor for the major surface antigens of  
RT plasmodium falciparum merozoites: studies at the genetic level."  
RL EMBO J. 4:3823-3829(1985).  
RN [2]  
RP REVISIONS, SEQUENCE FROM N.A.  
RA Fan W., Tolle R., Bujard H.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
CC (Potential).  
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42  
CC kDa and 19 kDa antigens which are the major surface antigens of  
CC merozoites. The maturation take place during schizont.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X03371; CAA27070.1; -;  
DR InterPro; IPR006209; EGF\_like.  
DR Pfam; PF00008; EGF; 1.  
KW Malaria; Merozoite; Polypeptide; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor.  
FT SIGNAL 1 19 POTENTIAL. MEROZOITE SURFACE PROTEIN 1.  
FT CHAIN 20 1630 MEROZOITE SURFACE PROTEIN 1.  
FT DOMAIN 67 84 TRIPEPTIDE SG(TP) REPEAT.  
FT TRANSMEM 1614 1630 MEMBRANE ANCHOR.  
FT CARBOHYD 97 97 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 755 755 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 759 759 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 774 774 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 835 835 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 911 911 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 955 955 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1156 1156 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1436 1436 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1517 1517 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEBC3CE0A46322 CRC64;  
Query Match 5.0%; Score 190.5; DB 1; Length 1630;  
Best Local Similarity 20.6%; Pred. No. 0.05;  
Matches 171; Conservative 138; Mismatches 255; Indels 265; Gaps 47;  
QY 5 ENRL-LNESSSSSGGLLYFSDINLFOAPMVVTSSTGDLSPSSSELENIPSENQVFSQA 63  
DB 918 ENILSLGKNKIYQELIGQKSSSE-NF-----YKILKDSPTFYNE 956  
QY 64 IWSGFIKVKKSDVEYTFATPSADNHTVMWDDQEVINKASNS-----NKIRLEKRLY- 114  
DB 957 SFTNFKVSKADD-----INSLNDESKKKLEEDINKLKTQLSFDLYNKYKLERLPD 1011  
QY 115 -----QIKTYQRENTEKGLDFKLYWTDQSNQKKEVISDNILQPLKQKSNRSKRS 168  
DB 1012 KKTGVGKYKQIKKLTLLKEQLESKL--NSLNNPKHVL--QNFSPVFNKKGAEIAETEN 1067  
QY 169 TSAGPTVPDRDNGDIPSLSEVGVTVVDVKNKRTFLSPWISNIHEKKGLTKY--KSSP-- 223  
DB 1068 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1094  
QY 224 ---EKWSTADSPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAYPIVHVD 265  
DB 1095 TLSEESITQEDNVALENFKVLSKLEGLKNDLNLNLEKKLSYLSGHLHLIA-----E 1147  
QY 266 MENILSKNEDQSTQNTDSTRITISKNTSRTHTSEVHGNVAHSAFFDGGSVAGFS 325  
DB 1148 LKEVIKKNK---YTGNSPSENNT-----DVNNALSEYKKFLPEGTDVATVVS 1191

QY 326 NSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARLNANRYVNTGTAPYVLPPT--- 381  
DB 1192 ESGSDTLEQSPKPKPASTHVGABS---NTITTSQ-NVDDEVDDVIIVPFGSEEDYDDL 1247  
QY 382 -SLVLGKNOTLATIDADENOLSQLAPNNYPSKNLAPIA---LNAQKFFSSTPTITMNYN 437  
DB 1248 QGVVTEAVTPSVID---NILSKI---ENEYEVLYLKPLAGVYRSLSKKOLENNVMTFVN 1301  
QY 438 -----QFLELEKTKQLRLDQV-YGNIAI-----YNFNGRVRVDTGSGNWSVL 481  
DB 1302 VKDILNSRFKNKFNKV-LESDLIPYKDLTSSNYVVKDPYKPLNKEKDKFLSSNYIK 1360  
QY 482 PQICEITARIIFNGKDLNLVERRIAAVNPSPDL-----ETTKPDM-TLKALKIAPGF 533  
DB 1361 DSID-----TDINFA-----NDVLGYKILSEKYKSLDKIKKVINDKQGE 1401  
QY 534 NEPN-----GNLOYQGGKITE-----FDNFDDQQTQSNITKNQJLAELNATNIYT 576  
DB 1402 NEKYLPPLANNIETLYKTVDNKIDLFVHLEAKVLYNTEK---SNVEVRIKELN--YLKT 1456  
QY 577 VLDKIKLNKAKN--ILIRDKRFHYDNN-----IAGV-ADESIVVKEAHREVINSSTEGLL 628  
DB 1457 IQDKLIADFKNNNFVGIAIDTSTYNNHLLTKELSTGMVFENLAKTVLSNLLDGNLQGG-M 1515  
QY 629 LNIDKD--IRKILSGYIVIEIETDEGLKVINDRYDMLNISLSRQDGKTFIDFKYNDKLP 686  
DB 1516 LNISQHCQVKK-----QCPQNSGCFRLHDE-----RECKCLLYKQEGDKC- 1557  
QY 687 LYSNPNYKVNVAVTYKENTIIINPSENG-----DTSNGIKKI 724  
DB 1558 --VENPNPTCN-----EN-----NGGCDADAKTEEDSGSNG-KKI 1590  
RESULT 4  
MSPI\_PLAFW STANDARD; PRT; 1639 AA.  
ID AC P04933;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)  
DE (Pwmsa) (P195).  
GN MSP-1.  
OS Plasmodium falciparum (isolate Wellcome).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5948;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86014355; PubMed=2995820;  
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,  
RA Nicholas S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,  
RA Freeman R.R.;  
RT "Primary structure of the precursor to the three major surface  
RT antigens of Plasmodium falciparum merozoites."  
RL Nature 317:270-273(1985).  
RN [2]  
RP REVISIONS  
RA Holder A.A.;  
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
CC (Potential).  
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42  
CC kDa and 19 kDa antigens which are the major surface antigens of  
CC merozoites. The maturation take place during schizont.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----



```

RX MEDLINE=89172073; PubMed=3148491;
RA Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppia S.H.,
RA Schmidt J.J.;
RA "Sequence and analysis of the DNA encoding protective antigen of
RT Bacillus anthracis.";
RL Gene 69:287-300(1988).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=Steine;
RC MEDLINE=99445483; PubMed=10515943;
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,
RA Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,
RA Martinez Y., Ricke D., Svensson R., Jackson P.J.;
RA "Sequence and organization of pXOI, the large Bacillus anthracis
RT plasmid harboring the Anthrax toxin genes.";
RL J. Bacteriol. 181:6509-6515(1999).
CC CC -I- SUBCELLULAR LOCATION: Membrane-bound (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
-----
DR ENBL; M22589; AAA22636.1; -.
DR ENBL; AF065404; AAD32415.1; -.
DR PIR; G59104; G59104.
DR PIR; I39933; I39933.
DR HSSP; PI3423; IACC.
KW Hypothetical protein; Plasmid; Transmembrane.
FT TRANSMEM 162 182
FT CONFLICT 93 93
FT CONFLICT 184 204
FT REF. 1)
FT KSCNCILIVVEVSQLMNSVFY -> NHVIVLISM (IN
FT REF. 1)
SQ SEQUENCE 204 AA; 23029 MW; E1657B23AE4273FD CRC64;
Query Match 6.2%; Score 235.5; DB 1; Length 204;
Best Local Similarity 34.6%; Pred. No. 1.7e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;
Qy 587 MNIIIDKRFHYDRNIAVGADSVVKEAHREVINGSTEGLLNIIDKDIEKILSGYIVEI 646
Db 1 MNIIIVRDP-YHYDNNNGNVGVDDSYLNKAYQKILWNSSGVSINLDEVDNQALSGYMLQI 59
Qy 647 EDTE-----GLKEVINDRYDMLINISLRQDGKTFIDFKKYNDKLPYISNP 693
Db 60 KPFSNHLTNSPVITLAGKDSGVGYLYRVLS-----DTGTGFDFNKFDENWASLV-DPG 112
Qy 694 YKNVTAVTKEN-TIINPSNGDTSTNGIKKILFSGYVEI 734
Db 113 DDVVYVAVKEDFNNAVTRDENGNTA-NKLXNTLVLSGKIKEI 153
RESULT 3
MSPI_PLAFK
ID MSPI_PLAFK STANDARD; PRT; 1630 AA.
AC P04932;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)
DE (PMWSA) (P190).
GN MSP-1.
OS Plasmodium falciparum (isolate KI / Thailand).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI TaxID=5939;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=86136024; PubMed=3004972;
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,
RA Stunnenberg H., Bujaard N.;

```

RL J. Biol. Chem. 266:15493-15497(1991).  
RN [6]  
RP CHARACTERIZATION.  
RC STRAIN=Sterne;  
RX MEDLINE=94327640; PubMed=8051159;  
RA Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;  
RT "Anthrax protective antigen forms oligomers during intoxication of  
RN mammalian cells.";  
RL J. Biol. Chem. 269:20607-20612(1994).  
RN [7]  
RP CHARACTERIZATION.  
RX MEDLINE=21129592; PubMed=11207581;  
RA Beauregard K.E., Collier R.J., Swanson J.A.;  
RC "Proteolytic activation of receptor-bound anthrax protective antigen  
RN on macrophages promotes its internalization.";  
RL Cell. Microbiol. 2:251-258(2000).  
RN [8]  
RP TOXIN REGULATION.  
RX STRAIN=Weybridge;  
RA Medline=94131936; PubMed=8300513;  
RC Koehler T.M., Dai Z., Kaufman-Yarbray M.;  
RA "Regulation of the Bacillus anthracis protective antigen gene: CO2 and  
RN a trans-acting element activate transcription from one of two  
RP promoters.";  
RL J. Bacteriol. 176:586-595(1994).  
RN [9]  
RP MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.  
RX STRAIN=Sterne;  
RA MEDLINE=95050722; PubMed=7961869;  
RC Singh Y., Klimpel K.R., Aroxa N., Sharma M., Leppla S.H.;  
RA "The chymotrypsin-sensitive site, PFD315, in anthrax toxin protective  
RN antigen is required for translocation of lethal factor.";  
RL J. Biol. Chem. 269:29039-29046(1994).  
RN [10]  
RP MUTAGENESIS OF DOMAIN 4 LOOPS.  
RX STRAIN=Sterne;  
RA MEDLINE=99185012; PubMed=10085028;  
RC Varughese M., Teixeira A.V., Liu S., Leppla S.H.;  
RA "Identification of a receptor-binding region within domain 4 of the  
RN protective antigen component of anthrax toxin.";  
RL Infect. Immun. 67:1860-1865(1999).  
RN [11]  
RP MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.  
RX STRAIN=Sterne;  
RA MEDLINE=21092804; PubMed=11178978;  
RC Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;  
RA "Trp 346 and Leu 352 residues in protective antigen are required for  
RN the expression of anthrax lethal toxin activity.";  
RL Biochem. Biophys. Res. Commun. 281:186-192(2001).  
RN [12]  
RP MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.  
RX STRAIN=Sterne;  
RA MEDLINE=21438996; PubMed=11554763;  
RC Ahuja N., Kumar P., Bhatnagar R.;  
RA "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are  
RN required for oligomerization of anthrax protective antigen.";  
RL Biochem. Biophys. Res. Commun. 287:542-549(2001).  
RN [13]  
RP MUTAGENESIS OF PRO-289.  
RX STRAIN=Sterne;  
RA MEDLINE=21253689; PubMed=11356563;  
RC Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;  
RA "Role of residues constituting the 2betaal strand of domain II in the  
RN biological activity of anthrax protective antigen.";  
RL FEMS Microbiol. Lett. 199:27-31(2001).  
RN [14]  
RP MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.  
RX MEDLINE=21125576; PubMed=1122612;  
RA Mogridge J., Moutreux W., Collier R.J.;  
RN "Involvement of domain 3 in oligomerization by the protective antigen  
RP moiety of anthrax toxin.";  
RL J. Bacteriol. 183:2111-2116(2001).  
RN [15]

RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.  
RX MEDLINE=21269403; PubMed=11113126;  
RA Sellman B.R., Nassi S., Collier R.J.;  
RN "Point mutations in anthrax protective antigen that block  
RP translocation.";  
RL J. Biol. Chem. 276:8371-8376(2001).  
RN [16]  
RP MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;  
RX ILE-239; TRP-255 AND PHE-265.  
RC STRAIN=Sterne;  
RA MEDLINE=22112896; PubMed=12117959;  
RC Chauhan V., Bhatnagar R.;  
RA "Identification of amino acid residues of anthrax protective antigen  
RN involved in binding with lethal factor.";  
RL Infect. Immun. 70:4477-4484(2002).  
RN [17]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RX MEDLINE=97192099; PubMed=9039918;  
RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;  
RN "Crystal structure of the anthrax toxin protective antigen.";  
RL Nature 385:833-838(1997).  
RN [18]  
RP REVIEW.  
RX MEDLINE=21428689; PubMed=11544370;  
RA Mock M., Fouet A.;  
RN "Anthrax.";  
RL Annu. Rev. Microbiol. 55:647-671(2001).  
CC -!- FUNCTION: One of the three proteins composing the anthrax toxin,  
CC the agent which infects many mammalian species and that may cause  
CC death. PA binds to a receptor (ATR) in sensitive eukaryotic  
CC cells, thereby facilitating the translocation of the enzymatic  
CC toxin components, edema factor and lethal factor, across the  
CC target cell membrane. PA associated with LF causes death when  
CC injected. PA associated with EF produces edema. PA induces  
CC immunity to infection with anthrax.  
CC -!- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a  
CC protective antigen (PA), a lethal factor (LF) and an edema factor  
CC (EF). None of these is toxic by itself. PA+LF forms the lethal  
CC toxin (Lctx); PA+EF forms the edema toxin (Edtx). PA-63 forms  
CC heptamers and this oligomerization is required for LF or EF  
CC binding. Once activated, at low pH, the heptamer undergoes  
CC conformational changes and converts from prepore to pore inserted  
CC in the membrane, forming cation-selective channels.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: The molecule is folded into four functional domains. Each  
CC domain is required for a particular step in the toxicity process.  
CC Domain 1 contains two calcium ions and the proteolytic activation  
CC site. Cleavage of the PA monomer releases the subdomain 1a, which  
CC is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is  
CC part of the remaining 63-kDa fragment (PA63) and contains the  
CC binding sites for LF and EF. Domain 2 is a beta-barrel core  
CC containing a large flexible loop that has been implicated in  
CC membrane insertion and pore formation. There is a chymotrypsin  
CC cleavage site in this loop that is required for toxicity. Domain 3  
CC has a hydrophobic patch thought to be involved in protein-protein  
CC interactions. Domain 4 appears to be a separate domain and shows  
CC limited contact with the other three domains; it would swing out  
CC of the way during membrane insertion. It is required for binding  
CC to the receptor; the small loop is involved in receptor  
CC recognition.  
CC -!- PTM: Proteolytic activation by furin or a furin-like protease  
CC cleaves the protein in two parts, PA-20 and PA-63; the latter is  
CC the mature protein. The cleavage occurs at the cell surface and  
CC probably in the serum of infected animals as well; both naive and  
CC cleaved PA are able to bind to the cell receptor. The release of  
CC PA20 from the remaining receptor-bound PA63 exposes the binding  
CC site for EF and LF, and promotes oligomerization and  
CC internalization of the protein.  
CC -!- MISCELLANEOUS: In ref.9 multiple mutagenesis experiments were  
CC performed that showed that the residues present in the small loop  
CC of domain 4, and not the ones in the large loop, are involved in  
CC receptor recognition.  
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 8.2539 Seconds  
(without alignments)  
4636.784 Million cell updates/sec

Title: US-09-848-909A-11  
Perfect score: 3774  
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILFSGKGVIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3745	99.2	764	1 PAG_BACAN	P13423 bacillus an
2	235.5	6.2	204	1 YPSI_BACAN	P13422 bacillus an
3	190.5	5.0	1630	1 MSP1_PLAFK	P04932 plasmodium
4	190.5	5.0	1639	1 MSP1_PLAFW	P04933 plasmodium
5	174.5	4.6	979	1 P115_MYCHR	P41508 mycoplasma
6	173	4.6	1658	1 YM67_YEAST	Q03661 saccharomyc
7	171	4.5	1000	1 S155_YEAST	P43612 saccharomyc
8	170.5	4.5	1230	1 SMC3_YEAST	P47037 saccharomyc
9	164.5	4.4	2334	1 WAPA_BACSU	Q07833 bacillus su
10	163	4.3	1385	1 FATI_SCHPO	O13735 schizosacch
11	162	4.3	1276	1 BXD_CLOBO	P19321 clostridium
12	160	4.2	1803	1 YJL3_YEAST	P47024 saccharomyc
13	159	4.2	1420	1 SRB3_YEAST	P38931 saccharomyc
14	158.5	4.2	1127	1 CNGA_HELPJ	Q921t1 helicobacte
15	158	4.2	1937	1 SPOR_SCHPO	Q14111 schizosacch
16	157.5	4.2	2116	1 MYS2_DICDI	P08799 dictyostell
17	157	4.2	1790	1 USO1_YEAST	P25386 saccharomyc
18	156.5	4.1	1487	1 MDS3_YEAST	P53094 saccharomyc
19	156	4.1	2867	1 RBP2_PLAVB	Q00799 plasmodium
20	155.5	4.1	1208	1 PCPI_SCHPO	Q92351 schizosacch
21	155.5	4.1	1460	1 N159_YEAST	P40477 saccharomyc
22	155	4.1	938	1 YMJ3_CABEL	P34487 caenorhabdi
23	154	4.1	1029	1 RIP3_RAT	Q9ere6 rattus norv
24	153.5	4.1	1233	1 YF16_YEAST	P43597 saccharomyc
25	152.5	4.0	1024	1 RIP3_MOUSE	P97434 mus musculu
26	152.5	4.0	1116	1 YK54_AQUAE	Q67838 aquifex aeo
27	152.5	4.0	1271	1 Y338_MYCGE	P47580 mycoplasma
28	152.5	4.0	2869	1 RBP1_PLAVB	Q00798 plasmodium
29	152	4.0	1882	1 Y462_MYCPN	P75109 mycoplasma
30	151.5	4.0	1397	1 CID_DROME	P19538 drosophila
31	150.5	4.0	1358	1 SIR4_YEAST	P11978 saccharomyc
32	150.5	4.0	1628	1 NAGH_CLOPE	P26831 clostridium
33	150	4.0	1037	1 KCC4_YEAST	P25389 saccharomyc

P53968 saccharomyc  
O14157 schizosacch  
P40468 saccharomyc  
P13611 homo sapien  
Q00496 clostridium  
Q92KWS helicobacte  
P16466 proteus mir  
P09959 saccharomyc  
P10844 clostridium  
Q03213 saccharomyc  
Q49413 mycoplasma  
P18494 saccharomyc

ALIGNMENTS

RESULT 1

ID	PAG_BACAN	STANDARD	PRT	764 AA.
AC	P13423; Q9FSR7; Q9RH69; Q9RQ2;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins			
DE	translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].			
GN	PAGA OR PAG OR PXOI-110.			
OS	Bacillus anthracis.			
OG	Plasmid pXOI.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
OX	NCBI_TaxID=1392;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=69172073; PubMed=3148491;			
RA	Welkos S.D., Lowe J.R., Eden-McCutchan F., Vockin M., Leppla S.H.,			
RA	Schmidt J.J.;			
RT	"Sequence and analysis of the DNA encoding protective antigen of			
RT	Bacillus anthracis.";			
RL	Gene 69:287-300(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=28, 33, BA1024, and BA1035;			
RX	MEDLINE=99214082; PubMed=10197996;			
RA	Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;			
RT	"Genetic diversity in the protective antigen gene of Bacillus			
RT	anthracis.";			
RL	J. Bacteriol. 181:2358-2362(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=V770-NP1-R / ATCC 14185;			
RX	MEDLINE=20359347; PubMed=10899854;			
RA	Cohen S., Mendelson I., Altboum Z., Kobiler D., Elhanany E., Bino T.,			
RA	Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,			
RA	Kronman C., Velan B., Shaffer A.;			
RT	"Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus			
RT	anthracis spore vaccines protect against anthrax.";			
RL	Infect. Immun. 68:4549-4558(2000).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sterner;			
RX	MEDLINE=99445483; PubMed=10515943;			
RA	Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,			
RA	Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,			
RA	Martinez Y., Rieke D., Svensson R., Jackson P.J.;			
RT	"Sequence and organization of pXOI, the large Bacillus anthracis			
RT	plasmid harboring the Anthrax toxin genes.";			
RL	J. Bacteriol. 181:6509-6515(1999).			
RN	[5]			
RP	DOMAINS.			
RX	MEDLINE=91332080; PubMed=1651334;			
RA	Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;			
RA	"The carboxyl-terminal end of protective antigen is required for			
RT	receptor binding and anthrax toxin activity.";			

Db 1051 NTKN-----TWPNYTLISELTSKEDSKNSITNSSKSDIET 1087

RESULT 15

T31105

hypothetical protein 2 - Haemophilus ducreyi

C:Species: Haemophilus ducreyi

C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999

C:Accession: T31105

R:Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.

J. Bacteriol. 180, 6013-6022, 1998

A:Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.

A:Reference number: Z20984; MUID:99030326; PMID:9811662

A:Accession: T31105

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4919 <WP>

A:Cross-references: EMBL:AF057696; NID:g3929021; PID:g3929023; PIDN:AAC79761.1

C:Genetics:

A:Gene: lrpA2

Query Match 4.7%; Score 178; DB 2; Length 4919;

Best Local Similarity 20.0%; Pred. No. 2.4;

Matches 182; Conservative 126; Mismatches 328; Indels 274; Gaps 41;

QY 6 NLLNESESSQGLGYFSDLNFOAPMVVTSSTGDL-SIPSSSELENIPSENQYF---Q 61

Db 154 NNSIKEGNSQLVLLGEN-KNLGSOAAKTIHQVIGDOESKISGGLVFGKADLIINP 212

QY 62 SAIWSGFIVKKSDEYTFATS-----DFKLYWTDSONK-----KEV 145

Db 213 NGVTLNGVKTINTDRFVASTSEVPEPHIKQLNVQRGKIIGKGVATNGLSHFVDVAKNI 272

QY 83 -----ADNRYVMVDQDVINKASNSKIRLEKRLGYOIKI-- 118

Db 273 EQQKGVSIKEDSKPAKLANVTFAAGNLTYDVNTRDV-NRNTNPKPITDKONIAISG 331

QY 119 -----CYQREN-----TEKGL-----DFKLYWTDSONK-----KEV 145

Db 332 ESAGSMYGRNIFIVTRGAGVNHQGVIFAEDDIIMILTDGNSRLNKVYADYRVVVGKDI 391

QY 146 ISDNLQLPELKQKSSNRKKSSTAGPTVPDRDNDGIPD-SLEVEGYTVDPKRTFLS 204

Db 392 ELANNQIHAHQQLILNATHGVKLDNGSSVINNNLIGISALNLTLENATVSANN---LS 447

QY 205 PMSNTHKKGKLTYSKSPKSTASDPYS-----PEKVTGRIDKNVSPEARHP 254

Db 448 FRYTNDTKLNLKSVS-----ARAADLQSGNLNLDKASVLAHKULTNISNDVSLNQSK 501

QY 255 LVAAYPVHVDMENIILSKNEQSTONTSETRTISKNTSTRTSTSEVHGNAEVEHAF 314

Db 502 LSA-----NNLKIKVRLNLNNSLSANNLTNTSNNITLKNK-----SKP 543

QY 315 DIGGSVAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTARLNANIRVYVGTGAPI 374

Db 544 TAGNMT---LNVTVNVTLNDSBELA-ANNLTLNVTXNVTLNDSKLSANKLDLV----- 594

QY 375 YNVLPTTSLVLGKQQLATIDADENQLS01-----LAPNNYPSKNLAPALNAQKF 427

Db 595 -----TDNVTLSKSTLSAGELTFKQKVVTLNDSLSAANN-----LSLNA----- 636

QY 428 SSTPITMYNQFLEKT--KQLRL---DTQVYGNIAFYNPENGRVVDVTGNSWSEVLP 482

Db 637 -SHNVTLNKKSLSAQKADIKAVNLTLNDTTE---LTAKNLDINSTTITNNGTIAGIFA 691

QY 483 QIQETTARIIFNGKDLNLYERRI-AAVNPSDPLEITKPDWTLKEALKIAPGENEENGNIQ 541

Db 692 NI---TTEKLNKKEKALILAEQNLNFTVNGSH---YENKGDIVSKDRTVTFKNS----- 741

QY 542 YQKGDITEFDNFDOQTSONIKNQLAELNATNI-----YTVLDKIKLNK----- 586

Db 742 -----DFTSGSKLVNAQNL-KVNVNFTISQGGDITLIGNVTLNASGTTFTNSGN 791

QY 587 ---MNLIRDKRFHYDRNNIAVGAD---ESVVK-----EAHREVINS-- 622

Db 792 LTTVKTLVDYGDIGNFTNKGKGLTVGEDLHKSKTKITNDGKGLSIKLNLSSEADFINNGT 851

QY 623 -----STEGLLNIDKDIRKILSGYVIEIEDTEGLKEVINDRYDM---LNISS--- 667

Db 852 LLGIEALKIATKGNFTNKEKAI--LASNGLDISVAEGKKTFNNGTIESGKLNLTNTGA 909

QY 668 -LRQDGKTFIDFKKYNDKLPYISN-----PNVKVNVYAVTKENTINPSENGDTSITNGI 721

Db 910 FLNVDMATRSFGLVLTITSTGVNSNNGTILISHERLN---ITSAANFTNES-NGTVMNSGL 965

QY 722 KKILIFSKKG 731

Db 966 LNII--AKQG 973

Search completed: May 3, 2004, 19:41:19

Job time : 16.0855 secs

A;Cross-references: GB:AE05176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146  
A;Experimental source: strain IL403  
C;Genetics: Ydgf

Query Match 4.8%; Score 179.5; DB 2; Length 1072;  
Best Local Similarity 19.4%; Pred. No. 0.21;  
Matches 160; Conservative 142; Mismatches 342; Indels 179; Gaps 30;

```
QY 2 VQENELLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSELEN----- 52
DB 300 VEQDVVASBESTQDANSALYPISEASVYDNLNLSISSLDSSISSTOTENSQSGASTA 359
QY 53 ---IPSENYFOSA---IWSGTFIKVKSDEYTFATSDADNHYVMVDQEVINKASNSNKI 106
DB 360 EISYDSENSLSLSSNQINSNSSEKDSQSSGLSSMSNEESHSNSNINETNSSEI 419
QY 107 R-----LEKRLYQIKIQRENPTKGLDFKL----YWTDSQNKKEVSSNLDLPK 157
DB 420 TNLPPNPTENSVDQTSSEASTNSNSISLSPSNISSTSDSESATNSDFSNAVA 479
QY 158 QKSSNRKKRSTAGTVPDRNDGI---POSLEVEGYTVVANKRTFLSPWISNIHEK 213
DB 480 NNSLASVNNSSSVLSTSTADNLGINQSGSDNLTKD--SSEISTSGAFLS---SNQTS 534
QY 214 KGLTKYKS-----TASDPISDFEKTGTRIDKNVSPPEARPLVAAP 261
DB 535 EASTNSNSSISLSPNISSTVLESTSSNSFNVAEAVNGLASVNNSSSVLSTSTA 594
QY 262 VHYDM-----ENT-----ILSKNEDOSTONTSETRTISKTSTRTSTSEVH 304
DB 595 DNEINQFGSDNLTKDSSEISTSGAFLSSNQTSSEASSNSMSINSPLSLSTNSESA 654
QY 305 GNAEVHASFPDGGSVAGFNSNSSTVAIDHSLSL--AGERTWETMGLNTADTARLNA 362
DB 655 TNO-----SNSSEATKVDNNSST---HSSNINLGSNDSDSDSDSDSNL-- 699
QY 363 NIRVNTGTAPIYVNLPTTSLVKGKQTLATIDADENOLSQ---ILAPNYYPSKULAP 418
DB 700 -----SSPNLETQTISSKSEFVNINSEPKVSSNSVQENSTDE 742
QY 419 IALNAOKFFSTPTMYNYOFLKTKQLRLDQVYGNIAFYN---FENGVRVDTGS 475
DB 743 MSTNPXSSISPTSTSSSQKESQN---LLNTTEGINPTFNSSSENSAASILT-- 797
QY 476 NWSEVLQIOETIARIF-----NGKDL-----NLVERRIAANVPSDPLETTKPD 522
DB 798 SYNNNSSESSETGLYISNEAQDNGSEISHLSPSSNENNVSISQSAILESSKSTN 857
QY 523 LK-EALKIATGFNEP-NGNLQYQKQDITEFDNFDDQTSQNIKNQLAELNATNIYTVLDK 580
DB 858 KRSSLSLSIINSTSHPNQEDNQSSD-----EVKSNNVBSILGQLNSISNKTWMS 909
QY 581 I---KLNAKNILIRDKRFHYDRNNTAVGADESVEKAEHREVINSGTEGLLINIDKIRK 637
DB 910 LTSQKLSVITLPSKSKVTNEKNENSTVSEKLIKTPQCN-DESONLGOITALDLSFNK 968
QY 638 ILSGYIVEIEDTEGLKEVIN-----DRYDMLNISSLRQD-GKTFIDFPKYNKPLPY 688
DB 969 -----EVETMEDSKTVPKVLNENGDNRQSNKSTIAKDKNKVFPKRSBENS 1019
QY 689 ISNPYKVNVAVTKENTINPENGDTSTNGKILIPSKG 731
DB 1020 -----DSDNNILKTVLLKKG 1036
```

RESULT 14

JC6009  
surface-located membrane protein lmp3 precursor - Mycoplasma hominis  
C;Species: Mycoplasma hominis  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C;Accession: JC6009  
R;Ladefoged, S.A.; Jensen, L.T.; Brock, B.; Birkelund, S.; Christiansen, G.

J. Bacteriol. 178, 2775-2784, 1996  
A;Title: Analysis of 0.5-kilobase-pair repeats in the Mycoplasma hominis lmp gene syse  
A;Reference number: JC6009; MUID:96213016; PMID:8631664  
A;Accession: JC6009  
A;Molecule type: DNA  
A;Residues: 1-1302 <LAD>  
A;Cross-references: EMBL:X95601; NID:g1197335; PIDN:CAA64858.1; PID:g1197336  
C;Genetics:  
A;Gene: lmp3  
A;Genetic code: SGC3  
C;Superfamily: surface-located membrane protein lmp3; tetratricopeptide repeat homology  
C;Keywords: duplication; membrane protein  
F;1-24/Domain: signal sequence #status predicted <SIG>  
F;25-1302/Domain: surface-located membrane protein lmp3 #status predicted <MAT>  
F;957-992/Domain: tetratricopeptide repeat homology <TT1>  
F;993-1026/Domain: tetratricopeptide repeat homology <TT2>  
F;1089-1120/Domain: tetratricopeptide repeat homology <TT3>  
F;1154-1190/Domain: tetratricopeptide repeat homology <TT4>

Query Match 4.7%; Score 178; DB 1; Length 1302;  
Best Local Similarity 18.0%; Pred. No. 0.33;  
Matches 138; Conservative 141; Mismatches 285; Indels 202; Gaps 28;

```
QY 1 EVKQENELLNSESQGLGYYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
DB 476 QVDEANKSIKEQNALIDKANTLLPOLNDSIVKAKSLNAEITNANKVQNQDNASM 535
QY 61 QSAIWSGFIKVKSDYTFATSDADNHYVMVDQEVINKASNSNKIRLEKGR--LYQIKI 118
DB 536 QSA-----KSSLDKVKTIQNLQTEFNKDKAKFKELEQTRKDIDNFLT 579
QY 119 QYQRENTEKGL-----DFKLYWTDSONKKEVSSDNLQLPELKQKSSNSR--KRS 168
DB 580 DDVKNKNYATLVKDLTNAKDKKSVTKSNKSEIIAAND-----ELKQALDKAKVADQI 635
QY 169 TSAGTVPDRNDGIPDSLEVEGYTV---DVKNKRTFLSPWISNIEKKGL----- 216
DB 636 DEANKSIKEQLSDSITNANQLLNKLVSDKDIOKAKTELSQEQSASQELNLANPTSMQS 695
QY 217 -----TKYKSSPEKWTASD--PYSDEPKVTGRIDKNVSPPEARPLVAAPVIVHDM 266
DB 696 AKESLDKAVTEITKKLETFNKKDKVFKFELEKTRKQIDEFINTKNTNP----- 743
QY 267 ENILSKNEDOSTONTSETRTISKTSTSTRTSTSEVHGNAEVHASFPDGGSVSAGFSN 326
DB 744 -----NYSTLISELTSKSDSKNSVTNSNKSIDI----- 771
QY 327 SNSSTVAIDHSLSLAG-ERTWAEETMGLNTADTAEALNANIRYVNTGTAPIYVNLPTTSLVL 385
DB 772 -ETANTELKQALAKANTDKAQADNLAKSTKE--QLNNSISSANTLLAKLTDKONTI--- 824
QY 386 GKNOTLATIDADENQLSQILAPNYYPSKULAPIALNAQ-----KPFSSPTPTMYNYQFL 440
DB 825 --QQAQTELEKEVQKAVASNN--TASMQSAKSLDAKVTETKKLE---TFNKDKDV 877
QY 441 ELETKQLRLDQVYGNIAFYNFENGVRVDTGSNNWSEVLQIOETIARIIIFNGKDLNL 500
DB 878 KFKELEQTRKDIDEF-----INTKNTPDYSTLISELTSK----- 912
QY 501 VERRIAANVPSDPLETTKPDMLKEALIAFGFNEPKNLQYQKQDITEFDNFDDQTSQ 560
DB 913 RDSKNSITNSNKSIDIETANTELKQALAKA-----NTDQKQAD-NLARSTKE 958
QY 561 NIKQQLAELNATNIYTVLDKIKLNAKNILIRDKRFHYDRNNTAVGAD---ESVVKBAHR 617
DB 959 QLNKSISAN-----TLAKLT-----DKDNTTQQAQTELEKEVQKAVQ 997
QY 618 EVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDR-YDMLNISSLRQDGKTFI 676
DB 998 AVAGNNTASM-----QSAKSLDAKVTET---TKLETFNKDKVFKFELEQTRKDIDEFI 1050
QY 677 DFKYNDKPLPYISNPYKVNVAVT-----KENTINPSENGDTST 718
```



Qy 385 LGKN-OTLATIDAD-ENQLSQI-----LAPN----NYYP---SKNLAPALNAQCKF 427  
Db 755 IDKSIETLTTEIKNSSENHIDEIKGQIDKLKVYFNKTMFNEDPKEIEKIENIVEKIDKKK 814  
Qy 428 S-----STPTMYNQFLEKTQKRLDQTDOVYGNIATYNFENGVRVDVTGSNWSE 479  
Db 815 NIYKEIXLLNEISKENDTSLKLNINLVSGKSLGNLFQQIDEEKKAETHIKAME 874  
Qy 480 V----LPQIOETTARIIFNGKDNL-----VERIAAVNPSDP-----LETT-----K 518  
Db 875 AVYDDLDNIIKKKSQEI---EKEMNNIMDIKMIDHKEMKALINTSHDDYKIIYHTSNKHBEK 931  
Qy 519 PDMTLCEALKIAFGFNEP--NGNLOYQGKDTIEDFPNDQOTSNIKQNLAELNATMIY 575  
Db 932 ISDIRKNSKLIQDFEESVINDIKELEKNVLE-----SQNNNTDINOYLKIE--NIY 984  
Qy 576 TVLDDKIKANAKMILIRDRF--HYDRNNIAGDAESVVKEAHREVINSSTGLLNIDK 633  
Db 985 NIL---KLAKIKIIDIQVKYEYTDIEKNN-----KKINAELSNS----- 1020  
Qy 634 DIRKILSGYIVETDEGLKE-----VINDRY---DMLNTSSLRQDGKTIDFKKYND 683  
Db 1021 --EKI----ITQKENSSLKECOSKITSIIDNVSECINKTNL-----KTVIVNEKKN- 1069  
Qy 684 KLPLYISN-PNYKNV-----YAVTKENTIINPSNGDTSNT-GTKKILIIFSK 730  
Db 1070 -INTYFKGAEBYQNVSLENFNNTEMADTKSQYLINIKKGNGTNTDYNIKELKEHKKK 1126

RESULT 10  
B89921  
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
C:Accession: B89921  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogucuma, A.; Mitutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; PMID:21311952; PMID:11418146  
A:Accession: B89921  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-6713 <R>  
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BAB42527.1; GSPDB:GN00149  
A:Experimental source: strain N315  
C:Genetics:  
A:Gene: ebha

[illegible]

286 QY TRII---SKNTSRTHTSEVHGNAEVAHSFFDGGVSA-----ITGQIDSATQVTGVQSVKDNATLNDNA 2771

2720 DB LATAKQOAKDALRQMTHLSDAQQS-----

338 QY L-----SLAGERTWAETMGLNTADTARLNNANIRYVNTGAPIVNVLPPTS-----LV 384

2772 DB MQLRNSIAKNDVKASQPVVDADTKQNA-----YNTAVTSAENIINATSQPTLDPASVT 2827

385 QY LGKNCOTLATIDABENQL--SOILAPNNYPSKNLAPIA--LNAOKFSSPTITMNYN-QF 439

2828 DB QAAANQ-----YNTAKTALNGAQNLANKKQETTANIRLSHLNNAKQKDLNTQVTPAPNIST 2883

440 QY LELEKTKQLRL-----DTDQV-----YGNIAI-----YNFEN 466

2884 DB VNOVKTKAEQLDQAMERLINGIQDKQVQKQSVNFTDADPEKQTAYNNAVTAENIINQAN 2943

467 QY GRURVDTGSWNSEVLPQIOE--TTARIIFNGKDLNLVRRI--AAVNPDSDELET-----516

2944 DB G-----TNANQSQVEAALSTVTTTKQALG-----DRKVTDAKNNANQTLSTLDNLNA 2999

517 QY -----TKPDMTILKEALKIAFGNEPNGLQ--YQKDIITEFDNFND-----555

2993 DB QKGAVTGNINQAHVAEVT--QA:QTAQELNTAMGNLKNLNDKDTLGSQNFADADPEK 3055

556 QY -----QQTSONIKQGLAEINATNIVTLDKIKLNKMKILLIRDKFPHYDRNNIAGADESV 611

3051 DB KNAYNEAVRNAENTLNKSTGNV--PKQOVEAAMN-----QVNTTKAALNGTQL 3098

612 QY VKEAHREVINSSTGL--LNLIDKDIKILSGYIEIEDTEGLKEVINDRYDMLNIS--667

3099 DB EXAKQHANTAI DGLSHLTNAQKEALKQLVQOQTTVAEACQNEKAN-----NVDAAM 3155

668 QY LQQ-----DGKTFIDFKYNDKPLIYSNPNKYNVY--AVTKENTLIINPSEN 713

3151 DB DKLRSQIADNATTKQNQNYTD-----ASPN-KKDAYNNAVTTTACGIIQDTTN 3196

RESULT 11

G90599

Hypothetical protein MYPU\_7030 [imported] - Mycoplasma pulmonis (strain UAB CT15)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

C:Accession: G90599

R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; M

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycop

A:Reference number: A9512; PMID:21267165; PMID:11353084

A:Accession: G90599

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-752 <KUR>

A:Cross-references: GB:AL445566; PID:g14090118; PIDN:CAC13876.1; GSPDB:GNC0153

C:Genetics:

A:Experimental source: strain UAB CT1P

A:Gene: MYPU\_7030

A:Genetic code: SGC3

Query Match 4.8%; Score 181; DB 2; Length 752;

Best Local Similarity 19.5%; Pred. No. 0.1;

Matches 161; Conservative 129; Mismatches 269; Indels 268; Gaps 433

QY 41 GDLSTPSSLENIIPSENQYFQSAIWGSFFIKVKSDEYTF-----ATSAOHVHTMWD 92

DB 20 GTISGYSIGISLPNESASL-----VKRAMELNFNLKQSVLNKSIIEILNQWD 69

QY 93 DOEVINKASN-----SNKIRLEKGR--LYQI---KIQ:QRENPTKGLDFKLYWTDQ 140

DB 70 ENKVSNI SANDFNKVQSKTPLENGEKITYSVLGKDIYFQIVNPSQKITSKVKITSSK 129

QY 141 NKEVITSSNLQLPELKKQSSN--SKKSTSTAGTVPDRND-----GIPDSLEVGSTV 194

DB 130 ISKDVYMD-----KQRLNDFAKRLVNFKSSASEKQSDIMASQFNKSKLEIKYL 192

3

QY 577 VLDKIKLNAKGN--ILIRDKRPHYDRNN-----IAGV-ADBSVVKAEHREIVNGSTEGLL 628  
Db 1466 IQDKLADFKKONNFVGIADLSTDYHNHLLTKFLSTGWFNFENLAKTVLSNLLDGNLQO-M 1524  
QY 629 LNKDKD--IRKILSGYIYEIEDTEGLKEVINDRYDMLNLSRLQDGTFFIDFKYNDKLP 686  
Db 1525 LNISQHCQVKK-----QCPNSGCFRHLDE-----RBECKLLNFKQEGDKC- 1566  
QY 687 LYISNPNKYVYATKENTINPENG-----DSTNGIKKI 724  
Db 1567 --VENPNPTCN-----EN-----NGCCDADAKCTEEDSGNG-KKI 1599  
RESULT 8  
B64635  
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: B64635  
R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 368, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: B64635  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2529 <TOM>  
A:Cross-references: GB:AE000602; GB:AE000511; NID:92314060; PIDN:AAD07969.1; PID:9231406  
Query Match 4.9%; Score 185.5; DB 2; Length 2529;  
Best Local Similarity 21.0%; Pred. No. 0.37;  
Matches 158; Conservative 94; Mismatches 261; Indels 241; Gaps 36;  
QY 123 ENTEGLDFKLYWTSQNK---KEVISSDNLQLPELKQKSSNRKSTAGTVPDR 178  
Db 729 QNPFAESVSGVWL--QKNTYNGKIYID-----PNSGQSGQNTLSTYANLF--- 779  
QY 179 DNDGIPDSLEVEGYVDVKNKRTFLSP---WT-----SNIHEKKGLTKYK 220  
Db 780 ---GRGFSVNIQNTLIGNTSVNSGLIWHGGFGYITGTFSAANIY---LTNNF 832  
QY 221 SPSKYST-----ASP-----PYSDFEKTGIDKNVSPEARHPLVAAPIV 262  
Db 833 KTGEGVNSDGGGANITFKASDNITWDGLNYDAETVTKMTQGS---OHSYATFDALN 889  
QY 263 HVDMEIILSKNEDQSTQNTDSETRTISKNTSTSRTHTEVHGNAEVHASFDDIGSVSA 322  
Db 890 NISVTNSSFS-----DMTWGRKFSFSAKNISFS-----NASF-----S 921  
QY 323 GFSNSNSSTVA---IDHSLAGERWAEWTMGNTADTARLANIRYVTGTAPIYVNL- 378  
Db 922 GFTNPGGSSVISANATNSLSFINSR-----LNGGAVYNLQANSIFPN-NTQAVFNVL 973  
QY 379 -----PTTSVLGKQNLATIDADENQLSOILAPNPNYPSKNIAPIAL--NAOKKFS 428  
Db 974 SRGTSFNATQLLGNFTLSS-----QSLANFGDTTLQNNANITLGNKSQAAPK 1025  
QY 429 STPTMNYNLFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNSNVEVLPIQET 488  
Db 1026 NS-LTLDDNNSLSDNGSVLNANNTSAFNQASNLINYGS-----QATF 1068  
QY 489 ARIIFNCKDLNL-VERRIAAVNPSPLETKPDMTLKEALKIAGFNEPNGLNOYQKDI 547  
Db 1069 NSLFFNGGTLNLASSKKNASNASFSNNTI---INLDDSVLSASNTSSLNANINFGASQ 1125  
QY 548 TEF-----DPNFQOQTSNIKQLA-----E 568  
Db 1126 ADFCGNTIIDTASPNFDSASSLNFNLNLTANGALNFNGYTPSLTKALMSVSGQFVLGNNG 1185  
QY 569 LNATNIYVLDKIKLNAKMLILDKRF-----HYDRNNIAV 605

Db 1186 INLSDI-NIFDNTKSVTYNILNAQKGTIGISGANGYEKILFYGMKIQNATYSDNNNIQT 1244  
QY 606 GA-----DESVVKEHR-----EVIN--SSTEGLLNIDKDI-----RKILSGYIVE 645  
Db 1245 WSPINPLNSQIIQESIKNGDLTIEVINPNPNSASNTIFNIAPELYNQASKQNPYISVD 1304  
QY 646 IEDTEGLKEVINDRYDMLNLSL-----RQDGTFFIDFKYNDKLPY---ISNPNY 694  
Db 1305 YSDNQA-----GTYYLTSNLIKGLFTPKGSQTPQAPGTYSFPFNOPLSSLNIYNGKFSSEN 1359  
QY 695 KVVYAVTKENTIN--PSENGDTSTNGIKKI 725  
Db 1360 KTLGLILSQNSATLKEMIESNQLDNITN-INEVL 1392  
RESULT 9  
T28576  
rhostry protein - Plasmodium yoelii (fragment)  
C:Species: Plasmodium yoelii  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T28576; A45521  
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.  
Mol. Biochem. Parasitol. 76, 329-332, 1996  
A:Title: Comparison of two members of a multigene family coding for high-molecular mass  
A:Reference number: Z20507; MUID:97077455; PMID:8920022  
A:Accession: T28576  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-2401 <SIN>  
A:Cross-references: EMBL:U36927; NID:91041784; PID:91041785; PIDN:AAB41263.1  
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-246, 1990  
A:Title: Identification of the gene for a Plasmodium yoelii rhostry protein. Multiple c  
A:Reference number: A45521; MUID:9101660; PMID:2270106  
A:Accession: A45521  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 2260-2401 <XEE>  
A:Cross-references: GB:M34281  
Query Match 4.9%; Score 184.5; DB 2; Length 2401;  
Best Local Similarity 21.6%; Pred. No. 0.38;  
Matches 181; Conservative 120; Mismatches 316; Indels 221; Gaps 43;  
QY 1 EVKQENLLNESESSQGLGYFSDLN-----FOAPMVVTSSTGDLSI 45  
Db 402 EYAKENVOLNRYKNLEIKKHNDQINIDNIKEAKQNYDQFKEMKTIIPNEMKYQK 461  
QY 46 PSSELENIPISENOYFQSAI--WSGFIKV---KKSDEYTFATSADNHYTMWVDDQEV---I 97  
Db 462 PSIBIKIMKDE---FLSKVKNKINDFKVYKEKVESEHNKFTLTNKTKEVSDBEIKKYE 518  
QY 98 NKASNSKIRLEKRLYQIKIQORENPTEKGLDFKLTDSQNKKEVISEDNLQLPK 157  
Db 519 NKFNDSKSLNETKK--SIEEYQNIINTLKKVDD---YIKVCLNTNELINTCHNQKQTLK 573  
QY 158 QKSNNSRK--KRSTSAGTVPDRDNDGIPD-----SLEVEGYVDVKNKRTFLSP 205  
Db 574 DKLQNKTIKETNSIDKIYTDKFNITDKKTELETKFTGLSLNNHESNNKELLTYFD 633  
QY 206 WISNHEKGLTKYKSPKSTASDPSYDFEKTGIDKNVSPEARHPLVAAPIVHVD 265  
Db 634 LKANLGNKNEMLYKQFNEK---EKAVEDIKKXNDINKIVSN-----IEIT 677  
QY 266 MENILSKNEDQSTQNTDSETRTISKNTSTSRHTSEVHGNAEVHASFDDIGSVSAGFS 325  
Db 678 IYTSIYNED-----TENE---IGKSIELLNTKVL-----KVKANVT 713  
QY 326 NSNSSTVAI--DHSLSLAGERTWAEWTMGNTADTARLANIRYVTGTAPIYVNLPTSLV 384  
Db 714 NLNEIKELKDYDFQDFGK-----EKNIKYPDEN--KIKNDITLQNK 754



A:Cross-references: GB:AL445566; PID:gl4090108; PIDN:CAC13866.1; GSPDB:GN00153  
A:Experimental source: strain UAB CTIP  
C:Genetics:  
A:Gene: MIPU\_6930  
A:Genetic code: SGC3

```
Query Match          5.1%; Score 191.5; DB 2; Length 1125;
Best Local Similarity 19.2%; Pred. No. 0.055;
Matches 166; Conservative 163; Mismatches 328; Indels 207; Gaps 40;

QY 3 KOENLLNESSESSGGLGYFSDNFQAPMVVTSSTGDISPSSELENIPISENQVFS 62
DB 120 ENDNSVNNKNSK-----NDEN-----LTLTKVRLGHVNV-LNQGNIPTKNLAISK 168
QY 63 AIWS-----GPIKV-----KXSDYTFATSDADNHVTMWVDDQ2-----VI 97
DB 169 VILHNKLDVLGLTEITNENGVTIINELNKSSDKWNVIVSLKRGTTGSSGQBEHVGI 228
QY 98 NKASNSKRLP-----KRLVQIKIOYORENPTKGL-----130
DB 229 YK-----ENKLTLSFDDKESKGFYENKLV-----DPTFKGQKIDVRPPFGVKSTKGN 282
QY 131 --DFKLYW--TDSQNKKE--VISSDNLQLPKQSSNRKRSKSTAGTVPDRDNGI 183
DB 283 KNDFTVVFHSDAPGVKEERGEISAKYSGQGHKEVAEALRTKEVMEYFDSI-----DGV 337
QY 184 PDSLEVEGYT-----VVKVKRTTSLP-W-INSIHEK---K 214
DB 338 NNELFTMGDTNKLGNKAEKAFPLQSGYKSLIKDVKNENATSLAORWGEYANHYDKIYK 397
QY 215 GLTKYKSSP--EKWSTASDPYSDFEKTGRIDK-NVSPPEARHP-----LVAAYPI 261
DB 398 GDUKVNSGFYDLWKYFDDNLLNKEEFKTVZEYRSSKKAKDKYKGEYSYVLHAI SDHTI 457
QY 262 VYVDMENIILSKNEDOSTONTSETRTISKVTSSTHSTSEVHGNAEVAHFDDIGGSVS 321
DB 458 VYTDL--ILLQKOD---QNKSENK--DENSDNSKQNDKPKTNQEQNTQDSSKKIS 509
QY 322 AGFSNSNSVVAIDHSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVLPPT 381
DB 510 DASQNSNTNEKQDKLDSODE---SKNNAIKSQNDQKDSNLSSKNDTQPSKSSPOI 566
QY 382 SLVLGKQNTLATIDADENQLSQILAPNYPYKSLAPIALNAQKFSPTITWYNQFLE 441
DB 567 NPNLENNQEI SHSNGENDSDSKQNTSNGRQTKN-----DLRSEKQKQLT-----611
QY 442 LEKTKQLRLDQVGNIAVYNGRVR---VDTGSNWSEVLPOIQTETARIIFNGKDL 498
DB 612 ---TKNPSNSNVETKNETQNNNSSTKDEIDTSA-----KTQDSTNSLNKEEKT 661
QY 499 NLVERRI-----AAVNPSPDLETTKPDWTLKEALKI--AFGNFENGNLQVQKQDITE 549
DB 662 NOVETKNTBSNNSNSTKQENSSTKKEBISKSEGNVANSNSTKNQENIDNKKEBISK 721
QY 550 PDRFDQOTSONIKNO-LAEINATNIYTVLDKIKLNAKKN-----TLIRDKRPHYDRNI 603
DB 722 SESVNNNSNTNQETPTETNSQNNVILGKPNQNSQNAIDAIVSAKKVIGYWNINE 781
QY 604 AVGADESIVKEAHEVINSSTEGLL-----LNIDKIRKILSGYVIEIDTEGLKEVINDR 659
DB 782 SVGKSSASKAFVAKVIDHNKLDIVGIGGLVHEETLTKI-----VEEMNKLSKSSDK 834
QY 660 YDMLNISLRDQGTTP-IDFKYNDKLPYISNPNKYVNTVATKENT---IINPSENGD 715
DB 835 W--VQVISEKKQGGFPVNLARYTG--VIYKEN---KFNIESFPKQNTNGHLYENQPN 897
QY 716 TSTNGIKKI-----LIFSCKG 731
DB 888 SSFNTSEKVSIVRPPFGIKFSTKG 911
```

RESULT 7  
S05603

major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (s  
N; Alternate names: gp195 surface antigen  
C; Species: Plasmodium falciparum  
C; Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 09-Jun-2000  
C; Accession: S05603; S04850  
R; Wyler, P. J.  
Submitted to the EMBL Data Library, April 1989  
A; Reference number: S05603  
A; Accession: S05603  
A; Molecule type: mRNA  
A; Residues: 1-1639 <MYL>  
A; Cross-references: EMBL:X15063; NID:g9896; PIDN:CAA33163.1; PID:g9897  
R; Wyler, P. J.  
Nucleic Acids Res. 17, 5401, 1989  
A; Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas  
A; Reference number: S04850; MUID:89345116; PMID:2668887  
A; Accession: S04850  
A; Molecule type: mRNA  
A; Residues: 1504-1639 <MYL2>  
A; Cross-references: EMBL:X15063  
C; Superfamily: major merozoite surface antigen  
C; Keywords: glycoprotein; merozoite; surface antigen  
F; 1-19/Domain: signal sequence #status predicted <SIG>  
F; 20-1639/Product: major merozoite surface antigen #status predicted <MAT>

```
Query Match          5.0%; Score 190.5; DB 2; Length 1639;
Best Local Similarity 20.6%; Pred. No. 0.11;
Matches 171; Conservative 138; Mismatches 255; Indels 265; Gaps 47;

QY 5 ENEL-LNESSESSGGLGYFSDNFQAPMVVTSSTGDISPSSELENIPISENQVFS 63
DB 927 ENILSGKNKIYQELIGQKSE-NF-----YEKIKDSDFYNE 965
QY 64 IWSGFIVKKSDEYTFATSDADNHVTMWVDDQEVINKASNS-----NKIRLEKGRLY- 114
DB 966 SPTNFVSKADD-----INSLNDESKRKKLEEDINKTKTLQSLFDLYNKYKILKLERLD 1020
QY 115 -----QIKYOQRENTEKGLDFKLYWDSQNKKEVISSDNLQLPKQSSNSRKR 168
DB 1021 KKKTGVYKMQIKKLTLLKEQLESKL--NSLNNPKHVL--QNTSVFNKKKEAIEIETEN 1076
QY 169 TSAGTVPDRDNDGIPDSLEVEGYTVDKNKRFLSPWISNIHEKKGLTKY---KQSP-- 223
DB 1077 T-----LENTKILLKHY-----KGLVYKNGESSPLK 1103
QY 224 ---EKWSTASDPYSDFE-----KVTGRIDKNVSPPEAR-----HPLVAAYPIVHVD 265
DB 1104 TLSSEIQTEDNYASLENFKVLSKLEKJKNLNLNLEKKLSYSSGLHLLIA-----E 1156
QY 266 MENIILSKNEDQSTONTSETRTISKVTSSTHSTSEVHGNAEVAHFDDIGGSVSAGFS 325
DB 1157 LKEVIKKNV---YTCNSPSENNT-----DVNNALESYKKEFLPEGTDAVTVS 1200
QY 326 NSNSTVAIDHSLSLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVLPPT---- 381
DB 1201 ESGSTLQSQPKKPPASTHVGAESE---NTITTSQ-NVDDEVDVIIIVPIFGESEEDVDD 1266
QY 382 -SLVLGKQNTLATIDADENQLSQILAPNYPYKSLAPIALNAQKFSSTPTITWYN 437
DB 1257 GQVTVGEAVTPEVID---NILSKI---ENEYEVLYLKLACGVRSLLKQLENNVTFTVN 1310
QY 438 -----QFLELEKTKQLRLDQV-YGNIAI-----YNFENGVRVVDGTGNWSEVL 481
DB 1311 VKDILNSRFNKRENFQNV-LESDLIPYKDLTSSNVVVKDPYKFLNKKERDKFLSSYNIK 1369
QY 482 POIQETTARIIFNGKOLNVERRIAANVPSDPL-----ETTKPDM-TLKEALKIAPGF 533
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKSDLSIKKYINDKQGE 1410
QY 534 NEPN-----GNLQYQKQITE-----FDNFQOTSONIKNLQELNATNIYT 576
DB 1411 NEKLPFLNNIETLYKTVNDKIDLFIHLEAKVLNITYEK---SNVEVKIKELN--YLKT 1465
```

QY 587 MNILIRKRFHYDRNNAIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVEI 646  
DB 1 MNILVRDP-YHYDNGNIVGVDDSVLKAYKQILNWSDDGVSLNDEVDNALSGYMLQI 59  
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDPKKYNKLPYISNPN 693  
DB 60 KPSPNHLTNSPVTTITLAGKDSGVGLYRVL-----DAGFLDPNKFENRSLV-DPG 112  
QY 694 KYVNVYAVTKEN-TIINPSENGDTSTNGIKKILIPSKGYEI 734  
DB 113 DDVYVYAVTKEDFNNAVTRDENGNI-NKUKNTLVLSGKIKEI 153

## RESULT 4

G59104  
hypotheical protein pXOI-111 - Bacillus anthracis virulence plasmid pXOI  
C:Species: Bacillus anthracis  
C:Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 11-May-2000  
C:Accession: G59104  
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler, J. Bacteriol. 181, 6509-6515, 1999  
A:Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbored by B. anthracis strain Sterne  
A:Reference number: A59091; MUID:99445483; PMID:10515943  
A:Accession: G59104  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-204 <OKI>  
A:Cross-references: GB:AF065404; NID:G4894216; PIDN:AA032415.1; PID:G4894327  
A:Experimental source: strain Sterne  
A>Note: similar to hypothetical protein in the protective antigen domain; ypa, plasmid pXOI  
C:Genetics:  
A:Gene: pXOI-111  
A:Genome: plasmid

Query Match 6.2%; Score 235.5; DB 2; Length 204;  
Best Local Similarity 34.6%; Pred. No. 2.6e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNILIRKRFHYDRNNAIAGADESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVEI 646  
DB 1 MNILVRDP-YHYDNGNIVGVDDSVLKAYKQILNWSDDGVSLNDEVDNALSGYMLQI 59  
QY 647 EDETE-----GLKEVINDRYDMLNLSLRQDGKTFIDPKKYNKLPYISNPN 693  
DB 60 KPSPNHLTNSPVTTITLAGKDSGVGLYRVL-----DAGFLDPNKFENRSLV-DPG 112  
QY 694 KYVNVYAVTKEN-TIINPSENGDTSTNGIKKILIPSKGYEI 734  
DB 113 DDVYVYAVTKEDFNNAVTRDENGNI-NKUKNTLVLSGKIKEI 153

## RESULT 5

F82885  
hypotheical protein U0482 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: F82885  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Casseil, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum. Alternate views of a min  
A:Reference number: A82870  
A:Accession: F82885  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-4688 <GLA>  
A:Cross-references: GB:AE002145; GB:AF222894; NID:G6899476; PIDN:AAF30894.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: U0482  
A:Genetic code: SGC3

Query Match 5.4%; Score 202; DB 2; Length 4688;  
Best Local Similarity 21.3%; Pred. No. 0.13;

Matches 174; Conservative 131; Mismatches 325; Indels 186; Gaps 40;  
QY 19 LLGYFSDLNQAPMVVTSSTTGDLSIPSELENIPISENQYFQSAIWSGFIKVKKSDYET 78  
DB 3699 LUDVYLD-NIHQNDITRKIPKDHV-SKEIIEINPVTMISKGNKWKSPPTTTANFEFK 3756  
QY 79 FATSADNHVMMVDDEVINKASNNKIRLEKGRLYQIKIQORENPTKGLDFKLYWTD 138  
DB 3757 IETQ-----DDNDVLNNIDATVKFKDEHNNIKQIVRIKEN-----ND 3795  
QY 139 SQNKKEVISSDNLQLPKOKSSN-----SRKXRSASAGPTVDPD----- 179  
DB 3796 WLIKQI-----DNLN-PETKYKLENTLSPLKTHNLVSINDKENISILITETGNPVLK 3851  
QY 180 ---NDGIPDSLEVEGYTDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWNSTASDYS 235  
DB 3852 IOTONDITNDTQOTINVTLSGVNSK-YNGRQIKVYKDNVNNVIYESS---LITLQKGD 3907  
QY 236 FEKVTRIGDKVNSPEARHPLVAAPYIVHYDMENILSKN-EDOSTONTSET-RTISKNT 293  
DB 3908 YQLLNLSNLSN-----REYRFEKIEIHNISNTNPFEDLEKUNGVSNTFITQKNT 3957  
QY 294 STSRHTS-EVHGNAEVSHPAS-----FDIGGSVSAGFS-----NSNS---STVAID 335  
DB 3958 TVQWNDSSATIVGTRGVNFNFKIKSEDKILENNQVVAWPAFAPKETIRDTNMLQYTRPLK 4017  
QY 336 HSLSLAGERTWATMGINT---ADTARLNANIRYNTGTAPIYVNLPTTSLVLGKQTLA 392  
DB 4018 DVTSDPKEGFWADLSNSVNFKEETTKLVKIQFVNKPKAKNNINNSNNVLDNTNSI 4077  
QY 393 ---TIDADENQLSQILAPNNYPSKNLAPALNAQK-FSSTPTITWY--NQFLELE 443  
DB 4078 NSNYEFTTKVGDHKLINITSNNVNTSQTINFTLSGVKSKVGVKKIKLSYKSNDSSES 4137  
QY 444 KTKQLRLDTPQVGNATATNFENG---VRVDTGSNWSEVLPQIQTETARIENGKD 497  
DB 4138 HTNEVLIESNKTQYNTLLNKLKENRTYTLIDVKLIDNNVSDPFKEGNTNSFITRTSA 4197  
QY 498 LNLVERRIAAVNSDPLETT-----KPDMLTKEA-----LKIAFGNEPENGLO 541  
DB 4198 INVLTIEISNRASNTWLKSTIIKINLNDPNDVLDRDQATIVYGNKNQKAMGFTTVSGNIK 4257  
QY 542 YQKQDTEFDNPDQOTSONIK-NQALAEUNATNYTVLDKIKUNAKONILIRDKRHYD- 599  
DB 4258 YLTATLVDLNFN-DKVNIVNISFNKPSAAEN-----IGIDKSNHII-----YNNDS 4304  
QY 600 ---RNNIATVGA---DESVVKEAAREVINSSTEGLLNIDKIRKILSGYIVEIETE 650  
DB 4305 IPKLEINNDIIVNGPINKELIVVKNANQK--NNIDVDLGLQINPKIAHNL-FLAKPKSTN 4361  
QY 651 GLKEVIND--RYDMLNLSL-RQDGKTFIDP-----KKY----- 681  
DB 4362 ---NDIETNVINGSSLVNNDGKTSIRFTLNLKANKLYSLVDVYVLVNNNSNIVE 4415  
QY 682 NDKLPYISNPNKYKVVATKENTINPSENGDTS 717  
DB 4416 SNKLP-KLANNINYQIK---INKSHITII--SKNGEWS 4445

## RESULT 6

E90598  
membrane nuclease, lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: E90598  
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moszer, I.  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul  
A:Reference number: A99512; MUID:21267165; PMID:11353084  
A:Accession: E90598  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-1125 <KUR>

```
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVMTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120
DB 90 QSAIWSGFIKVKKSDEYTFATSDADNHVMTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 149
QY 121 QRENTEKGLDFKLWYTDSONKKEVYSSDNLQLPKQKSSNSRKRKSTAGPTVPDRDN 180
DB 150 QRENTEKGLDFKLWYTDSONKKEVYSSDNLQLPKQKSSNSRKRKSTAGPTVPDRDN 209
QY 181 DGIPDSLEVEGYTVQKNTKRTFLSPWISNIHEKGLTKYKSPKWSASTDPYSDFEKVT 240
DB 210 DGIPDSLEVEGYTVQKNTKRTFLSPWISNIHEKGLTKYKSPKWSASTDPYSDFEKVT 269
QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 270 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 329
QY 301 SEVHGNAEVAHSPFDIGGVSAGFSNSSTVAIDHSLSLAGERTWAEATMGLNTADTARL 360
DB 330 SEVHGNAEVAHSPFDIGGVSAGFSNSSTVAIDHSLSLAGERTWAEATMGLNTADTARL 389
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQSLIAPNNYPSKNLAPIA 420
DB 390 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQSLIAPNNYPSKNLAPIA 449
QY 421 LNAQKFSSTPTMTNQNQLEKTKQLRLDQVYGNIAVYVNFENGVRVDTGNSKSEV 480
DB 450 LNAQKFSSTPTMTNQNQLEKTKQLRLDQVYGNIAVYVNFENGVRVDTGNSKSEV 509
QY 481 LPQIQETTARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGL 540
DB 510 LPQIQETTARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGL 569
QY 541 QYQKDIITFDNFDOOTSONIKQALAEALNATNYVLDKIKNAKKNILIRDKFHFYDR 600
DB 570 QYQKDIITFDNFDOOTSONIKQALAEALNATNYVLDKIKNAKKNILIRDKFHFYDR 629
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYVIEIDTEGLKEVINDRY 660
DB 630 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDKIRKILSGYVIEIDTEGLKEVINDRY 689
QY 661 DMLNISLRODGTFTDFKXNDKPLIYSNPNYKVNVAVTKNTIINPSENGSTSTNG 720
DB 690 DMLNISLRODGTFTDFKXNDKPLIYSNPNYKVNVAVTKNTIINPSENGSTSTNG 749
QY 721 IKKILIFSKGYBIG 735
DB 750 IKKILIFSKGYBIG 764

RESULT 2
I40862
iota toxin component Ib - Clostridium perfringens
C:Species: Clostridium perfringens
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999
C:Accession: I40862; S42774
R:Perelle, S.; Gibert, M.; Boquet, P.; Popoff, M.R.
Infect. Immun. 61, 5147-5156, 1993
A:Title: Characterization of Clostridium perfringens iota-toxin genes and expression in
A:Reference number: I40861; MUID:94041637; PMID:8225592
A:Accession: I40862
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-875 <RES>
A:Cross-references: EMBL:X73562; NID:g929031; PIDN:CAA51960.1; PID:g414655

Query Match 23.1%; Score 871; DB 2; Length 875;
Best Local Similarity 31.1%; Pred. No. 1.5e-36;
Matches 252; Conservative 131; Mismatches 278; Indels 150; Gaps 31;

QY 1 EVQENPLLNESSSSGLGYFSDNFCAPMVVTSSTTGSLSPSSLENIIPSE-NOY 59
DB 35 DTNQKEBITNLTSSNGLMGYFADEHFHFXDLEMAPIKNGDLKFEKKVYDKLTEDNS 94
```

```
QY 60 FOSAIWSGFIKVKKSDEYTFATSDADNHVMTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 119
DB 95 IKSIRWTGRIIPSEDEGYILSTDR-NDVLQINAKGDIKAK--TLKVNKKQAYNIRIE 150
QY 120 YQREN-----PTEKGLDFKLWYTDSONKKEVYSSDNLQLPKQKSSNSRKRKST 170
DB 151 IQKNLGSIDNLSVP-----KLW-ELUNGKTVIPEENLFFRDISKIDEND----- 195
QY 171 AGPTVP-----DRNDGIPDSLEVEGYTVQKNTKRTFLSPWISNIH 211
DB 196 --PFIENNFFDVRPFSAAWEDELDTDNDNIPDAVEKNGYTI---KDSIAVKWDSGA 249
QY 212 EKGLTKYKSSPEKWSASTDPYSDFEKVTGRDKNVSEARHPLVAAYPIVHVDMENIIL 271
DB 250 E-QYKYYVSSYLESTAGDPYTDYQKASGSDKAILEARDPLVAAYPVGVGEMENLII 308
QY 272 SKNEDOSTQNTDSETRTISKNTSTRTSEVHGNAEVAHSPFDIGGVSAGFSNSNST 331
DB 309 STNEHASS---DQKTVSRATNSKTANTV-----GVGISAGYQNGFTCN 351
QY 332 VAIDHS-----LSLAGERTWAEATMGLNTADTARLNAIRYVNTGTAPIYVNLPTTSL 383
DB 352 ITTSYSHTTDNSTPAVDNGESNMTGLSINKGESAYINANVRYNTGTAPMYKVFTTNL 411
QY 384 VLGNQTLATIDADENQSLIAPNNYPSKNLAPIALNAQKFSSTPTMTNQNQLELE 443
DB 412 VL-DGETLATIKADNQIGNNLSPNETYPKGLSPLALNTMQFNARLIPINYDQKKLD 470
QY 444 KTKQLRLDQVYGNIAVYVNFENGVRVDTGNSKSEVLPQIQETTARIIFNGKDLNVER 503
DB 471 SGQIKLETTQVSGNTGTKN-SQGI-I-TEGNSWSNYISQIDSVSASIIL-TGSGTFR 527
QY 504 RIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGLQY-QGKDITE--FDNFQDQTSQ 560
DB 528 RVAAKQCGNPEDT-PEITIGEAIKAFSATK-NGELLYFNGIPIIDESCVELIFDDNTSE 585
QY 561 NIKNQALAEALNATNYVLDKIKNAKKNILIRDKRF--HYDR-NNIAVGADESVMKEAHR 617
DB 586 IIEQLKYLDDKKIYVNV---KLERGMNIIKVPYSYFTNFDEYNNFP--ASWSNIDTKNQ 639
QY 618 EVINSSTEGI-----LLNIDKDKIRKILSGY-----IVEIEDTEGLKEVIND 658
DB 640 DGIQSVANKLSGTEKIIIPMSKLPKYRVFSGYKDPSTSNISITVNIKSKQKTDYLP 699
QY 659 RYDMLNIS-----SLRQDQGTFTDFKXND--KLPLYISNPNYKV----- 696
DB 700 EKDYTKFSYEFETTKDSSDIEITLTSSGVIFLDNLSITELNSTPEILKEPEIKVPSDOB 759
QY 697 -----NVAVTKNTIINPSENGDTSTNGI 721
DB 760 ILDAHNKYADIKLDT-----NTGNTYIDGI 785

RESULT 3
I39933
cryptic protein - Bacillus anthracis
C:Species: Bacillus anthracis
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999
C:Accession: I39933
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthracis
A:Reference number: I39933; MUID:89172073; PMID:3148491
A:Accession: I39933
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-192 <RES>
A:Cross-references: GB:M22589; NID:g143280; PIDN:AAA22636.1; PID:g143281

Query Match 6.2%; Score 235.5; DB 2; Length 192;
Best Local Similarity 34.6%; Pred. No. 2.4e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 Seconds  
(without alignments)  
5403.004 Million cell updates/sec

Title: US-09-848-909A-11

Perfect score: 3774

Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKKILFSGKGVIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

P.R. 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3748	99.3	764	2 I39934	protective antigen
2	871	23.1	875	2 I40862	iota toxin compene
3	235.5	6.2	192	2 I39933	cryptic protein -
4	235.5	6.2	204	2 G59104	hypothetical prote
5	202	5.4	4688	2 F82885	hypothetical prote
6	191.5	5.1	1125	2 E92598	membrane nucleas
7	190.5	5.0	1639	2 S05603	major merozoite su
8	185.5	4.9	2529	2 B64635	toxin-like outer m
9	184.5	4.9	2401	2 T28676	rhostry protein -
10	182	4.8	6713	2 B89921	hypothetical prote
11	181	4.8	752	2 G90599	hypothetical prote
12	181	4.8	2269	2 T28677	rhostry protein -
13	179.5	4.8	1072	2 A86827	hypothetical prote
14	178	4.7	1302	1 JC6009	surface-located me
15	178	4.7	4919	2 T31105	hypothetical prote
16	177.5	4.7	979	2 JQ0894	Pil5 protein - Myc
17	176	4.7	1227	2 C97033	uncharacterized pr
18	175	4.6	1631	1 SAZOK1	major merozoite su
19	174	4.6	1635	2 A10452	hemolysin [importe
20	173.5	4.6	2399	2 H71579	toxin-like outer m
21	173	4.6	1193	2 S68218	botulinum neurotox
22	173	4.6	1658	2 S55101	hypothetical prote
23	173	4.6	2178	2 S55805	alpha-toxin - Clos
24	170.5	4.5	786	2 T18469	hypothetical prote
25	170.5	4.5	821	2 S67087	hypothetical prote
26	170.5	4.5	1230	2 S56850	SMC1 protein homol
27	170.5	4.5	2340	2 F71704	cell surface antig
28	170	4.5	4152	2 T31102	filamentous hemaag
29	170	4.5	5005	2 F82884	hypothetical prote

## RESULT 1

I39934  
Protective antigen precursor - Bacillus anthracis plasmid  
C/Species: Bacillus anthracis  
C/Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 01-Dec-2000  
C/Accession: I39934; S69160; F59104  
R/Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.  
Gene 69, 287-300, 1988  
A/Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr  
A/Reference number: I39933; MUID:89172073; PMID:3148491  
A/Accession: I39934  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-764 <RES>  
A/Cross-references: GB:M22589; NID:G143280; PIDN:AAA22637.1; PID:G143282  
R/Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klimpel, K.R.; Birch, N.P.; Loh, Y.P.  
Arch. Biochem. Biophys. 316, 5-13, 1995  
A/Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SF  
A/Reference number: S69160; MUID:95142670; PMID:7840657  
A/Accession: S69160  
A/Molecule type: protein  
A/Residues: 197-202 <FRI>  
J/Okinaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle  
R. Bacteriol. 181, 6509-6515, 1999  
A/Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbor  
A/Reference number: A59091; MUID:99445483; PMID:10515943  
A/Accession: F59104  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-313, 'O', 315-764 <OKI>  
A/Cross-references: GB:AF065404; NID:G4894216; PIDN:AAD32414.1; PID:G4894326  
A/Experimental source: strain Sterne  
A/Note: similar to anthrax toxin moiety, protective antigen, paga formerly pag, plasmid  
C/Genetics:  
A/Gene: pXO1-110  
A/Genome: plasmid  
C/Function:  
A/Description: three component exotoxin; protective antigen binds to receptors on the s  
y active components edema factor or lethal factor; the complex is internalized by recep  
C/Keywords: exotoxin.  
F/1-29/Domain: signal  
F/130-196/Domain: propeptide #status predicted <PRO>  
F/197-202/Product: protective antigen #status experimental <MAT>

Query Match 99.3%; Score 3748; DB 2; Length 764;

Best Local Similarity 99.5%; Pred. No. 2.3e-182;

Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYQF 60

DB 30 EVKQENRLNSESQGLLYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENYQF 89

181 DGIIPDSLEVEGYVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKMWSTASDPYSDFEKVT 240  
181 DGIIPDSLEVEGYVDVKNKRTFLSPWISNIEHKKGLTKYKSSPEKMWSTASDPYSDFEKVT 240  
241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTWTGLTADTARL 360  
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTWTGLTADTARL 360  
361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
421 LNAQKXFSSTPTIMYNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480  
421 LNAQKXFSSTPTIMYNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480  
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540  
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540  
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILIRKRFHYDR 600  
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILIRKRFHYDR 600  
601 NNIAGADESVVKEAHEVINSSTEGLLINIDKDIRKILSGYIVEIETDEGLKEVINDRY 660  
601 NNIAGADESVVKEAHEVINSSTEGLLINIDKDIRKILSGYIVEIETDEGLKEVINDRY 660  
661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
721 IKKILIFSCKGYEIG 735  
721 IKKILIFSCKGYEIG 735

Search completed: May 3, 2004, 19:57:42  
Job time : 177.546 secs

241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTWTGLTADTARL 360  
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWAGTWTGLTADTARL 360  
361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
421 LNAQKXFSSTPTIMYNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480  
421 LNAQKXFSSTPTIMYNYNQFLEKTKQLRLDQVYGNIAATYVNGRVRVDTGNSWSEV 480  
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540  
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540  
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILIRKRFHYDR 600  
541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLILIRKRFHYDR 600  
601 NNIAGADESVVKEAHEVINSSTEGLLINIDKDIRKILSGYIVEIETDEGLKEVINDRY 660  
601 NNIAGADESVVKEAHEVINSSTEGLLINIDKDIRKILSGYIVEIETDEGLKEVINDRY 660  
661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
721 IKKILIFSCKGYEIG 735  
721 IKKILIFSCKGYEIG 735

RESULT 15  
PCT-US03-35733-6  
; Sequence 6, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-6

Query Match 99.3%; Score 3749; DB 1; Length 735;  
Best Local Similarity 99.5%; Pred. No. 2.4e-296;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 EVKQENRLLNESSESSQGLGYFSDINFPQPMVVTSTTGDLSPSSSELENIPSENQYF 60  
1 EVKQENRLLNESSESSQGLGYFSDINFPQPMVVTSTTGDLSPSSSELENIPSENQYF 60  
61 QSAIWSGFIKVKSDVETATSDNHNVTMVDQEVINKASNSNKIRLEKGRLYQIKIY 120  
61 QSAIWSGFIKVKSDVETATSDNHNVTMVDQEVINKASNSNKIRLEKGRLYQIKIY 120  
121 QRENPTKGLDPKLYWTDSONKKEVSSDNLQJPELKQKSSNRKRSSTAGTVPDRDN 180  
121 QRENPTKGLDPKLYWTDSONKKEVSSDNLQJPELKQKSSNRKRSSTAGTVPDRDN 180

361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPPSKNLAPIA 420  
361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420  
421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480  
421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480  
481 LPOIQTETARIIFNGKQNLNVERIAA VNPSPDLETTKPDMLKEALKIAFGFNEPNGNL 540  
481 LPOIQTETARIIFNGKQNLNVERIAA VNPSPDLETTKPDMLKEALKIAFGFNEPNGNL 540  
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNKNILIRDKGFHYDR 600  
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNKNILIRDKGFHYDR 600  
601 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660  
601 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660  
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
721 IKKILIFSKKGYEIG 735  
721 IKKILIFSKKGYEIG 735  
RESULT 13  
US-09-848-909A-7  
; Sequence 7, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-7  
Query Match 99.4%; Score 3750; DB 23; Length 735;  
Best Local Similarity 99.5%; Pred. No. 2e-296;  
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNNKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQJPELKQKSSNRKKKSTAGPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQJPELKQKSSNRKKKSTAGPTVPDRDN 180  
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
DB 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNBOQSTONTSETRTISKNTSRTHT 300  
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNBOQSTONTSETRTISKNTSRTHT 300

301 SEVHGNAEVHASFDFDIGSVSAGFSNSNSTVAIDHLSLAGERTWAGTGLNTADTARL 360  
301 SEVHGNAEVHASFDFDIGSVSAGFSNSNSTVAIDHLSLAGERTWAGTGLNTADTARL 360  
361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPPSKNLAPIA 420  
361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420  
421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480  
421 LNAQKFSSTPTIMYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480  
481 LPOIQTETARIIFNGKQNLNVERIAA VNPSPDLETTKPDMLKEALKIAFGFNEPNGNL 540  
481 LPOIQTETARIIFNGKQNLNVERIAA VNPSPDLETTKPDMLKEALKIAFGFNEPNGNL 540  
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNKNILIRDKGFHYDR 600  
541 QYQCKDITEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNKNILIRDKGFHYDR 600  
601 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660  
601 NNIAVGADESUVKEAHRVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660  
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
661 DMLNSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
721 IKKILIFSKKGYEIG 735  
721 IKKILIFSKKGYEIG 735  
RESULT 14  
PCT-US03-35733-4  
; Sequence 4, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-4  
Query Match 99.3%; Score 3749; DB 1; Length 735;  
Best Local Similarity 99.5%; Pred. No. 2.4e-296;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNNKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDEYTFATSADNHVTMVDQEVINKASNNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQJPELKQKSSNRKKKSTAGPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQJPELKQKSSNRKKKSTAGPTVPDRDN 180  
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
DB 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240

Db 481 LPOIETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGFNEPENG 540  
Qy 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAXNNILIRKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAXNNILIRKRFHYDR 600  
Qy 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEEDTEGLKEVINDRY 660  
Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEEDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRQDGKTFIDFKKNDKPLIYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKNDKPLIYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

## RESULT 11

US-09-848-909A-13  
; Sequence 13, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-13

Query Match 99.4%; Score 3753; DB 23; Length 735;  
Best Local Similarity 99.6%; Pred. No. 1.1e-296;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKSDSEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRN 180  
Qy 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240  
Db 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVIGKNOTLATIADENQLSQIILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVIGKNOTLATIADENQLSQIILAPNNYPSKNLAPIA 420

Qy 421 LNAQKFSSTPITWNYNOFLEKTKQLRLDQVYGNIAYNFENGVRVDTGNSHSEV 480  
Db 421 LNAQKFSSTPITWNYNOFLEKTKQLRLDQVYGNIAYNFENGVRVDTGNSHSEV 480  
Qy 481 LPOIETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGFNEPENG 540  
Db 481 LPOIETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGFNEPENG 540  
Qy 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAXNNILIRKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAXNNILIRKRFHYDR 600  
Qy 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEEDTEGLKEVINDRY 660  
Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIIEEDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRQDGKTFIDFKKNDKPLIYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKNDKPLIYISNPYKVNVAVTYKNTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

## RESULT 12

PCT-US03-35733-7  
; Sequence 7, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-7

Query Match 99.4%; Score 3750; DB 1; Length 735;  
Best Local Similarity 99.5%; Pred. No. 2e-296;  
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKSDSEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPDLKQKSSNSRKRSTSGPTVPDRN 180  
Qy 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240  
Db 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSSTASDPYDFEKT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300  
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Db 601 NNTAVGADBSVWKEAREVINSSTEGLLNIDKDIRKILSGYVIEDETEGLKEVINDRY 660  
Qy 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Db 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Qy 721 IKKILIFSKKGYSIG 735  
Db 721 IKKILIFSKKGYSIG 735

RESULT 9  
US-09-848-909A-8  
; Sequence 8, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-8

Query Match 99.5%; Score 3754; DB 23; Length 735;  
Best Local Similarity 99.6%; Pred. No. 9.5e-297;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLLGYFFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSQGLLGYFFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKSDSYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKRSTAGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKRSTAGPTVPDRDN 180  
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240  
Qy 241 GRIDKXVSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300  
Db 241 GRIDKXVSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300  
Qy 301 SEVHGNAEVSAPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360  
Db 301 SEVHGNAEVSAPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360  
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWEV 480  
Db 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWEV 480  
Qy 481 LPOIQTETARIIPNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPNGNL 540  
Db 481 LPOIQTETARIIPNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPNGNL 540  
Qy 541 QYQKDIITBDFNFDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600

Db 541 QYQKDIITBDFNFDQTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
Qy 601 NNTAVGADBSVWKEAREVINSSTEGLLNIDKDIRKILSGYVIEDETEGLKEVINDRY 660  
Db 601 NNTAVGADBSVWKEAREVINSSTEGLLNIDKDIRKILSGYVIEDETEGLKEVINDRY 660  
Qy 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Db 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720  
Qy 721 IKKILIFSKKGYSIG 735  
Db 721 IKKILIFSKKGYSIG 735

RESULT 10  
PCT-US03-35733-13  
; Sequence 13, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-13

Query Match 99.4%; Score 3753; DB 1; Length 735;  
Best Local Similarity 99.6%; Pred. No. 1.1e-296;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLLGYFFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSQGLLGYFFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKSDSYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKRSTAGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKKRSTAGPTVPDRDN 180  
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240  
Qy 241 GRIDKXVSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300  
Db 241 GRIDKXVSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRTH 300  
Qy 301 SEVHGNAEVSAPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360  
Db 301 SEVHGNAEVSAPFDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLTADTARL 360  
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWEV 480  
Db 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDVTGSNWEV 480  
Qy 481 LPOIQTETARIIPNGKDLNVERRIAANVPSDPLETTKPDMLTKEALKIAFGNEPNGNL 540



## RESULT 7

US-09-848-909A-2  
; Sequence 2, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-2

Query Match 99.5%; Score 3755; DB 23; Length 735;  
Best Local Similarity 99.6%; Pred. No. 7.9e-297;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	EVKQENRLNSES	SSQGLGYFSDLN	FQAPMVVTS	TTGDLSP	SSSELENIP	SENQYF	60																																																	
Db	1	EVKQENRLNSES	SSQGLGYFSDLN	FQAPMVVTS	TTGDLSP	SSSELENIP	SENQYF	60																																																	
Qy	61	QSAIWSGFI	KVKKSDEY	TATFATSAD	NHVTMW	DDQEV	INKASNSKIR	LEKGLYQIKIY 120																																																	
Db	61	QSAIWSGFI	KVKKSDEY	TATFATSAD	NHVTMW	DDQEV	INKASNSKIR	LEKGLYQIKIY 120																																																	
Qy	121	QRENTEKGLD	FKLYWTD	SONKEVI	SSDN	LQPEL	KQKSSNR	KRSTAGTVPDRDN 180																																																	
Db	121	QRENTEKGLD	FKLYWTD	SONKEVI	SSDN	LQPEL	KQKSSNR	KRSTAGTVPDRDN 180																																																	
Qy	181	DGIPDSLE	VEGYTV	VYKNT	FLSPWIS	NIHEK	GLTKYK	SSPEKWS	TASDPYDFEKT 240																																																
Db	181	DGIPDSLE	VEGYTV	VYKNT	FLSPWIS	NIHEK	GLTKYK	SSPEKWS	TASDPYDFEKT 240																																																
Qy	241	GRIDKNV	SPEARH	PLVA	APIV	HVD	MENI	ILSKNED	QSTNTDSTRIS	KNTS	SRHT 300																																														
Db	241	GRIDKNV	SPEARH	PLVA	APIV	HVD	MENI	ILSKNED	QSTNTDSTRIS	KNTS	SRHT 300																																														
Qy	301	SEVHGNAE	VHAS	FFD	IGGS	V	SAG	F	SNS	STVA	IDH	SL	S	L	A	G	E	R	T	W	A	E	T	M	G	L	N	T	A	D	T	A	R	L 360																							
Db	301	SEVHGNAE	VHAS	FFD	IGGS	V	SAG	F	SNS	STVA	IDH	SL	S	L	A	G	E	R	T	W	A	E	T	M	G	L	N	T	A	D	T	A	R	L 360																							
Qy	361	NANIRVNT	GTAPI	YNNV	PT	SLV	L	G	K	Q	T	L	A	T	I	D	A	D	E	N	Q	L	S	Q	I	L	A	P	N	N	Y	P	S	K	N	L	A	P	I	A 420																	
Db	361	NANIRVNT	GTAPI	YNNV	PT	SLV	L	G	K	Q	T	L	A	T	I	D	A	D	E	N	Q	L	S	Q	I	L	A	P	N	N	Y	P	S	K	N	L	A	P	I	A 420																	
Qy	421	LNAQK	F	S	T	P	I	T	M	N	Y	N	Q	L	E	K	T	Q	L	R	D	T	D	Q	V	Y	G	N	I	A	T	Y	N	F	E	N	G	R	V	R	V	D	T	G	S	N	W	S	E	V 480							
Db	421	LNAQK	F	S	T	P	I	T	M	N	Y	N	Q	L	E	K	T	Q	L	R	D	T	D	Q	V	Y	G	N	I	A	T	Y	N	F	E	N	G	R	V	R	V	D	T	G	S	N	W	S	E	V 480							
Qy	481	LPOI	Q	E	T	T	A	R	I	I	F	N	G	K	D	L	N	L	V	E	R	R	I	A	A	V	N	P	S	D	P	L	E	T	T	K	P	D	M	T	L	K	E	A	L	K	I	A	F	G	F	N	E	P	N	G	L 540
Db	481	LPOI	Q	E	T	T	A	R	I	I	F	N	G	K	D	L	N	L	V	E	R	R	I	A	A	V	N	P	S	D	P	L	E	T	T	K	P	D	M	T	L	K	E	A	L	K	I	A	F	G	F	N	E	P	N	G	L 540
Qy	541	QYQ	K	D	I	T	E	F	D	F	N	D	Q	O	S	T	S	Q	N	I	K	N	Q	L	A	E	L	N	A	T	N	I	Y	T	V	L	D	K	I	K	N	A	K	N	I	L	I	R	D	K	R	F	H	Y	D	R 600	
Db	541	QYQ	K	D	I	T	E	F	D	F	N	D	Q	O	S	T	S	Q	N	I	K	N	Q	L	A	E	L	N	A	T	N	I	Y	T	V	L	D	K	I	K	N	A	K	N	I	L	I	R	D	K	R	F	H	Y	D	R 600	
Qy	601	NNI	A	V	G	A	D	E	S	V	V	K	E	A	R	E	V	I	N	S	T	E	G	L	L	N	I	D	K	I	R	K	I	L	S	G	I	V	E	I	E	D	E	G	L	K	E	V	I	N	D	R	Y 660				
Db	601	NNI	A	V	G	A	D	E	S	V	V	K	E	A	R	E	V	I	N	S	T	E	G	L	L	N	I	D	K	I	R	K	I	L	S	G	I	V	E	I	E	D	E	G	L	K	E	V	I	N	D	R	Y 660				
Qy	661	DM	L	N	I	S	S	L	R	Q	D	G	K	T	F	I	D	F	K	K	Y	N	D	K	P	L	I	Y	I	S	N	P	N	Y	K	N	V	A	T	K	E	N	T	I	I	N	P	S	E	N	G	D	T	S	T	N	G 720
Db	661	DM	L	N	I	S	S	L	R	Q	D	G	K	T	F	I	D	F	K	K	Y	N	D	K	P	L	I	Y	I	S	N	P	N	Y	K	N	V	A	T	K	E	N	T	I	I	N	P	S	E	N	G	D	T	S	T	N	G 720

## RESULT 8

PCT-US03-35733-8  
; Sequence 8, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-8

Query Match 99.5%; Score 3754; DB 1; Length 735;  
Best Local Similarity 99.6%; Pred. No. 9.5e-297;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	EVKQENRLNSES	SSQGLGYFSDLN	FQAPMVVTS	TTGDLSP	SSSELENIP	SENQYF	60																																																	
Db	1	EVKQENRLNSES	SSQGLGYFSDLN	FQAPMVVTS	TTGDLSP	SSSELENIP	SENQYF	60																																																	
Qy	61	QSAIWSGFI	KVKKSDEY	TATFATSAD	NHVTMW	DDQEV	INKASNSKIR	LEKGLYQIKIY 120																																																	
Db	61	QSAIWSGFI	KVKKSDEY	TATFATSAD	NHVTMW	DDQEV	INKASNSKIR	LEKGLYQIKIY 120																																																	
Qy	121	QRENTEKGLD	FKLYWTD	SONKEVI	SSDN	LQPEL	KQKSSNR	KRSTAGTVPDRDN 180																																																	
Db	121	QRENTEKGLD	FKLYWTD	SONKEVI	SSDN	LQPEL	KQKSSNR	KRSTAGTVPDRDN 180																																																	
Qy	181	DGIPDSLE	VEGYTV	VYKNT	FLSPWIS	NIHEK	GLTKYK	SSPEKWS	TASDPYDFEKT 240																																																
Db	181	DGIPDSLE	VEGYTV	VYKNT	FLSPWIS	NIHEK	GLTKYK	SSPEKWS	TASDPYDFEKT 240																																																
Qy	241	GRIDKNV	SPEARH	PLVA	APIV	HVD	MENI	ILSKNED	QSTNTDSTRIS	KNTS	SRHT 300																																														
Db	241	GRIDKNV	SPEARH	PLVA	APIV	HVD	MENI	ILSKNED	QSTNTDSTRIS	KNTS	SRHT 300																																														
Qy	301	SEVHGNAE	VHAS	FFD	IGGS	V	SAG	F	SNS	STVA	IDH	SL	S	L	A	G	E	R	T	W	A	E	T	M	G	L	N	T	A	D	T	A	R	L 360																							
Db	301	SEVHGNAE	VHAS	FFD	IGGS	V	SAG	F	SNS	STVA	IDH	SL	S	L	A	G	E	R	T	W	A	E	T	M	G	L	N	T	A	D	T	A	R	L 360																							
Qy	361	NANIRVNT	GTAPI	YNNV	PT	SLV	L	G	K	Q	T	L	A	T	I	D	A	D	E	N	Q	L	S	Q	I	L	A	P	N	N	Y	P	S	K	N	L	A	P	I	A 420																	
Db	361	NANIRVNT	GTAPI	YNNV	PT	SLV	L	G	K	Q	T	L	A	T	I	D	A	D	E	N	Q	L	S	Q	I	L	A	P	N	N	Y	P	S	K	N	L	A	P	I	A 420																	
Qy	421	LNAQK	F	S	T	P	I	T	M	N	Y	N	Q	L	E	K	T	Q	L	R	D	T	D	Q	V	Y	G	N	I	A	T	Y	N	F	E	N	G	R	V	R	V	D	T	G	S	N	W	S	E	V 480							
Db	421	LNAQK	F	S	T	P	I	T	M	N	Y	N	Q	L	E	K	T	Q	L	R	D	T	D	Q	V	Y	G	N	I	A	T	Y	N	F	E	N	G	R	V	R	V	D	T	G	S	N	W	S	E	V 480							
Qy	481	LPOI	Q	E	T	T	A	R	I	I	F	N	G	K	D	L	N	L	V	E	R	R	I	A	A	V	N	P	S	D	P	L	E	T	T	K	P	D	M	T	L	K	E	A	L	K	I	A	F	G	F	N	E	P	N	G	L 540
Db	481	LPOI	Q	E	T	T	A	R	I	I	F	N	G	K	D	L	N	L	V	E	R	R	I	A	A	V	N	P	S	D	P	L	E	T	T	K	P	D	M	T	L	K	E	A	L	K	I	A	F	G	F	N	E	P	N	G	L 540
Qy	541	QYQ	K	D	I	T	E	F	D	F	N	D	Q	O	S	T	S	Q	N	I	K	N	Q	L	A	E	L	N	A	T	N	I	Y	T	V	L	D	K	I	K	N	A	K	N	I	L	I	R	D	K	R	F	H	Y	D	R 600	
Db	541	QYQ	K	D	I	T	E	F	D	F	N	D	Q	O	S	T	S	Q	N	I	K	N	Q	L	A	E	L	N	A	T	N	I	Y	T	V	L	D	K	I	K	N	A	K	N	I	L	I	R	D	K	R	F	H	Y	D	R 600	
Qy	601	NNI	A	V	G	A	D	E	S	V	V	K	E	A	R	E	V	I	N	S	T	E	G	L	L	N	I	D	K	I	R	K	I	L	S	G	I	V	E	I	E	D	E	G	L	K	E	V	I	N	D	R	Y 660				
Db	601	NNI	A	V	G	A	D	E	S	V	V	K	E	A	R	E	V	I	N	S	T	E	G	L	L	N	I	D	K	I	R	K	I	L	S	G	I	V	E	I	E	D	E	G	L	K	E	V	I	N	D	R	Y 660				
Qy	661	DM	L	N	I	S	S	L	R	Q	D	G	K	T	F	I	D	F	K	K	Y	N	D	K	P	L	I	Y	I	S	N	P	N	Y	K	N	V	A	T	K	E	N	T	I	I	N	P	S	E	N	G	D	T	S	T	N	G 720
Db	661	DM	L	N	I	S	S	L	R	Q	D	G	K	T	F	I	D	F	K	K	Y	N	D	K	P	L	I	Y	I	S	N	P	N	Y	K	N	V	A	T	K	E	N	T	I	I	N	P	S	E	N	G	D	T	S	T	N	G 720

; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof  
; FILE REFERENCE: PV595FCT  
; CURRENT APPLICATION NUMBER: PCT/US03/19786  
; CURRENT FILING DATE: 2003-06-25  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Mature PA sequence including an ETB signal sequence  
PCT-US03-19786-4

Query Match 99.6%; Score 3758; DB 1; Length 735;  
Best Local Similarity 99.6%; Pred. No. 4,5e-297;  
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY	1	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTS	TGDLSPSELENIPSENOYF	60
DB	1	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTS	TGDLSPSELENIPSENOYF	60
QY	61	QSAIWSGFIKVKSDYEYTFATSADNHVTWVDDQEV	INKASNKIRLEKGRLYQIKIY	120
DB	61	QSAIWSGFIKVKSDYEYTFATSADNHVTWVDDQEV	INKASNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLWTDPSQNKKEVISSDNLQLPEL	KOKSNRKRSTSGPTVPDRDN	180
DB	121	QRENPTKGLDFKLWTDPSQNKKEVISSDNLQLPEL	KOKSNRKRSTSGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGL	TKYKSSPEKWSASTASDPSDFEKT	240
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGL	TKYKSSPEKWSASTASDPSDFEKT	240
QY	241	GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNE	DQSTQNTDSETRTISKNTSRTHT	300
DB	241	GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNE	DQSTQNTDSETRTISKNTSRTHT	300
QY	301	SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHS	LSLAGERTWAETMGLNTADTARL	360
DB	301	SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHS	LSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATID	ADENQLSQILAPNNYPSKNLAPIA	420
DB	361	NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATID	ADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIA	TYNFENGVRVDTGSNWSEV	480
DB	421	LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIA	TYNFENGVRVDTGSNWSEV	480
QY	481	LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETT	KPDMLKEALKIATGFNPNGL	540
DB	481	LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETT	KPDMLKEALKIATGFNPNGL	540
QY	541	QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVL	DKIKLNKNNILIRDKRPHYDR	600
DB	541	QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVL	DKIKLNKNNILIRDKRPHYDR	600
QY	601	NNIAGADESVVKEAAREVINSSTGLLNIDKDKIRK	ILSGYIVIEIDTEGLKEVINDRY	660
DB	601	NNIAGADESVVKEAAREVINSSTGLLNIDKDKIRK	ILSGYIVIEIDTEGLKEVINDRY	660
QY	661	DMLNISLRODQKTFIDFKKYNKDKPLYSNPNYKVN	YAVTKENTIINPSENGDTSNG	720
DB	661	DMLNISLRODQKTFIDFKKYNKDKPLYSNPNYKVN	YAVTKENTIINPSENGDTSNG	720
QY	721	IKKILIFSKKGYEIG	735	
DB	721	IKKILIFSKKGYEIG	735	

RESULT 6  
PCT-US03-35733-2

; Sequence 2, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-2

QY	1	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTS	TGDLSPSELENIPSENOYF	60
DB	1	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTS	TGDLSPSELENIPSENOYF	60
QY	61	QSAIWSGFIKVKSDYEYTFATSADNHVTWVDDQEV	INKASNKIRLEKGRLYQIKIY	120
DB	61	QSAIWSGFIKVKSDYEYTFATSADNHVTWVDDQEV	INKASNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLWTDPSQNKKEVISSDNLQLPEL	KOKSNRKRSTSGPTVPDRDN	180
DB	121	QRENPTKGLDFKLWTDPSQNKKEVISSDNLQLPEL	KOKSNRKRSTSGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGL	TKYKSSPEKWSASTASDPSDFEKT	240
DB	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGL	TKYKSSPEKWSASTASDPSDFEKT	240
QY	241	GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNE	DQSTQNTDSETRTISKNTSRTHT	300
DB	241	GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNE	DQSTQNTDSETRTISKNTSRTHT	300
QY	301	SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHS	LSLAGERTWAETMGLNTADTARL	360
DB	301	SEVHGNAEVSASFDDIGGSVSAGFSNSSTVAIDHS	LSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATID	ADENQLSQILAPNNYPSKNLAPIA	420
DB	361	NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATID	ADENQLSQILAPNNYPSKNLAPIA	420
QY	421	LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIA	TYNFENGVRVDTGSNWSEV	480
DB	421	LNAQKFSSTPITMNYNQFLEKTKQLRLDQVYGNIA	TYNFENGVRVDTGSNWSEV	480
QY	481	LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETT	KPDMLKEALKIATGFNPNGL	540
DB	481	LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETT	KPDMLKEALKIATGFNPNGL	540
QY	541	QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVL	DKIKLNKNNILIRDKRPHYDR	600
DB	541	QYQKDIITEFDNFDOQTSQNKQLAELNATNIYTVL	DKIKLNKNNILIRDKRPHYDR	600
QY	601	NNIAGADESVVKEAAREVINSSTGLLNIDKDKIRK	ILSGYIVIEIDTEGLKEVINDRY	660
DB	601	NNIAGADESVVKEAAREVINSSTGLLNIDKDKIRK	ILSGYIVIEIDTEGLKEVINDRY	660
QY	661	DMLNISLRODQKTFIDFKKYNKDKPLYSNPNYKVN	YAVTKENTIINPSENGDTSNG	720
DB	661	DMLNISLRODQKTFIDFKKYNKDKPLYSNPNYKVN	YAVTKENTIINPSENGDTSNG	720
QY	721	IKKILIFSKKGYEIG	735	
DB	721	IKKILIFSKKGYEIG	735	

SEQ ID NO 10  
LENGTH: 735  
TYPE: PRT  
ORGANISM: Bacillus anthracis  
PCT-US03-35733-10

Query Match 99.7%; Score 3761; DB 1; Length 735;  
Best Local Similarity 99.7%; Pred. No. 2.5e-297;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	EVKQENLLNESESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
DB	1	EVKQENLLNESESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
QY	61	QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQDEVINKASNKIRLEKGRLYQIKIY	120
DB	61	QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQDEVINKASNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKRGSTAGPTVPDRDN	180
DB	121	QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKRGSTAGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTVDKVKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYVT	240
DB	181	DGIPDSLEVEGYTVDKVKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYVT	240
QY	241	GRIDKQVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT	300
DB	241	GRIDKQVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT	300
QY	301	SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
DB	301	SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQSLQILAPNNYPSKNLAPIA	420
DB	361	NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQSLQILAPNNYPSKNLAPIA	420
QY	421	LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV	480
DB	421	LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV	480
QY	481	LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMLKEALKIAGFNEPNGNL	540
DB	481	LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMLKEALKIAGFNEPNGNL	540
QY	541	OYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNIILIRDKRFHYDR	600
DB	541	OYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNIILIRDKRFHYDR	600
QY	601	NNIAGADESVVKEAHREVINSGTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY	660
DB	601	NNIAGADESVVKEAHREVINSGTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY	660
QY	661	DMNLSSLRQDGKTFIDFKKNDKPLIYSNPNYKVVAVTKENTIINPSENGDTSTNG	720
DB	661	DMNLSSLRQDGKTFIDFKKNDKPLIYSNPNYKVVAVTKENTIINPSENGDTSTNG	720
QY	721	IKKILIFSKKGYEIG 735	
DB	721	IKKILIFSKKGYEIG 735	

RESULT 4  
US-09-848-909A-10  
Sequence 10, Application US/09848909A  
GENERAL INFORMATION:  
APPLICANT: Sellman, R. John  
TITLE OF INVENTION: Compounds and Methods for the Treatment  
TITLE OF INVENTION: and Prevention of Bacterial Infection  
FILE REFERENCE: 00742/060002  
CURRENT APPLICATION NUMBER: US/09/848,909A

CURRENT FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: US 60/201,800  
PRIOR FILING DATE: 2000-05-04  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 735  
TYPE: PRT  
ORGANISM: Bacillus anthracis  
US-09-848-909A-10

Query Match 99.7%; Score 3761; DB 23; Length 735;  
Best Local Similarity 99.7%; Pred. No. 2.5e-297;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	EVKQENLLNESESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
DB	1	EVKQENLLNESESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF	60
QY	61	QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQDEVINKASNKIRLEKGRLYQIKIY	120
DB	61	QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQDEVINKASNKIRLEKGRLYQIKIY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKRGSTAGPTVPDRDN	180
DB	121	QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPKQKSSNRKRGSTAGPTVPDRDN	180
QY	181	DGIPDSLEVEGYTVDKVKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYVT	240
DB	181	DGIPDSLEVEGYTVDKVKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYVT	240
QY	241	GRIDKQVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT	300
DB	241	GRIDKQVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT	300
QY	301	SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
DB	301	SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQSLQILAPNNYPSKNLAPIA	420
DB	361	NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQSLQILAPNNYPSKNLAPIA	420
QY	421	LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV	480
DB	421	LNAQKFFSTPTIMYVNTQFLELEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV	480
QY	481	LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMLKEALKIAGFNEPNGNL	540
DB	481	LPQIQETTARIIFNGKDLNVERRIAANVPSPDPLETTKPDMLKEALKIAGFNEPNGNL	540
QY	541	OYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNIILIRDKRFHYDR	600
DB	541	OYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKKNIILIRDKRFHYDR	600
QY	601	NNIAGADESVVKEAHREVINSGTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY	660
DB	601	NNIAGADESVVKEAHREVINSGTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY	660
QY	661	DMNLSSLRQDGKTFIDFKKNDKPLIYSNPNYKVVAVTKENTIINPSENGDTSTNG	720
DB	661	DMNLSSLRQDGKTFIDFKKNDKPLIYSNPNYKVVAVTKENTIINPSENGDTSTNG	720
QY	721	IKKILIFSKKGYEIG 735	
DB	721	IKKILIFSKKGYEIG 735	

RESULT 5  
PCT-US03-19786-4  
Sequence 4, Application PC/TUS0319786  
GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc.

US-09-848-909A-11

Query Match 100.0%; Score 3774; DB 23; Length 735;  
Best Local Similarity 100.0%; Pred. No. 2.2e-298;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60

QY 61 QSAIWSGFIKVKKSDYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180  
DB 121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180

QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240  
DB 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQIILAPNNYPSKNLAPIA 420  
DB 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQIILAPNNYPSKNLAPIA 420

QY 421 LNAQKFSSTPTIMVNYQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480  
DB 421 LNAQKFSSTPTIMVNYQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

QY 541 QYQKDIETEPDFNFDQSTQNIKNQLAELNATNYTVLDKIKLNKAKNILLIRKRFHYDR 600  
DB 541 QYQKDIETEPDFNFDQSTQNIKNQLAELNATNYTVLDKIKLNKAKNILLIRKRFHYDR 600

QY 601 NNIAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
DB 601 NNIAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660

QY 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 3  
US-09-848-909A-11  
; Sequence 11, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis

US-09-848-909A-11

Query Match 100.0%; Score 3774; DB 23; Length 735;  
Best Local Similarity 100.0%; Pred. No. 2.2e-298;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60

QY 61 QSAIWSGFIKVKKSDYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDYFATSAADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180  
DB 121 QRENTEKGLDPKLYWTDSONKKEVIVSSDNQLPELKQKSSNSRKRSTSGPTVPDRN 180

QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240  
DB 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQIILAPNNYPSKNLAPIA 420  
DB 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQIILAPNNYPSKNLAPIA 420

QY 421 LNAQKFSSTPTIMVNYQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480  
DB 421 LNAQKFSSTPTIMVNYQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGSNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540

QY 541 QYQKDIETEPDFNFDQSTQNIKNQLAELNATNYTVLDKIKLNKAKNILLIRKRFHYDR 600  
DB 541 QYQKDIETEPDFNFDQSTQNIKNQLAELNATNYTVLDKIKLNKAKNILLIRKRFHYDR 600

QY 601 NNIAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
DB 601 NNIAVGADESVMKEAHREVINSTEGLLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660

QY 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 3  
PCT-US03-35733-10  
; Sequence 10, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSEQ for Windows Version 4.0

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds  
(without alignments)  
4086.665 Million cell updates/sec

Title: US-09-848-909A-11

Perfect score: 3774

Sequence: 1 EVKQENRLNSESSESSQGLL.....TSTNGIKILLIFSKKGHGIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/pctus COMB.pcp.\*  
2: /cgn2\_6/ptodata/2/paa/US06 COMB.pcp.\*  
3: /cgn2\_6/ptodata/2/paa/US07 COMB.pcp.\*  
4: /cgn2\_6/ptodata/2/paa/US08 COMB.pcp.\*  
5: /cgn2\_6/ptodata/2/paa/US081 COMB.pcp.\*  
6: /cgn2\_6/ptodata/2/paa/US082 COMB.pcp.\*  
7: /cgn2\_6/ptodata/2/paa/US083 COMB.pcp.\*  
8: /cgn2\_6/ptodata/2/paa/US084 COMB.pcp.\*  
9: /cgn2\_6/ptodata/2/paa/US085 COMB.pcp.\*  
10: /cgn2\_6/ptodata/2/paa/US086 COMB.pcp.\*  
11: /cgn2\_6/ptodata/2/paa/US087 COMB.pcp.\*  
12: /cgn2\_6/ptodata/2/paa/US088 COMB.pcp.\*  
13: /cgn2\_6/ptodata/2/paa/US089 COMB.pcp.\*  
14: /cgn2\_6/ptodata/2/paa/US090 COMB.pcp.\*  
15: /cgn2\_6/ptodata/2/paa/US091 COMB.pcp.\*  
16: /cgn2\_6/ptodata/2/paa/US092 COMB.pcp.\*  
17: /cgn2\_6/ptodata/2/paa/US093 COMB.pcp.\*  
18: /cgn2\_6/ptodata/2/paa/US094 COMB.pcp.\*  
19: /cgn2\_6/ptodata/2/paa/US095 COMB.pcp.\*  
20: /cgn2\_6/ptodata/2/paa/US096 COMB.pcp.\*  
21: /cgn2\_6/ptodata/2/paa/US097A COMB.pcp.\*  
22: /cgn2\_6/ptodata/2/paa/US097B COMB.pcp.\*  
23: /cgn2\_6/ptodata/2/paa/US098 COMB.pcp.\*  
24: /cgn2\_6/ptodata/2/paa/US099A COMB.pcp.\*  
25: /cgn2\_6/ptodata/2/paa/US099B COMB.pcp.\*  
26: /cgn2\_6/ptodata/2/paa/US100 COMB.pcp.\*  
27: /cgn2\_6/ptodata/2/paa/US101 COMB.pcp.\*  
28: /cgn2\_6/ptodata/2/paa/US102 COMB.pcp.\*  
29: /cgn2\_6/ptodata/2/paa/US103 COMB.pcp.\*  
30: /cgn2\_6/ptodata/2/paa/US104 COMB.pcp.\*  
31: /cgn2\_6/ptodata/2/paa/US106 COMB.pcp.\*  
32: /cgn2\_6/ptodata/2/paa/US107 COMB.pcp.\*  
33: /cgn2\_6/ptodata/2/paa/US06 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match Length	ID	Description
------------	-------	--------------	----	-------------

	1	3774	100.0	735	1	PCT-US03-35733-11	Sequence 11, Appl
	2	3774	100.0	735	23	US-09-848-909A-11	Sequence 11, Appl
	3	3761	99.7	735	1	PCT-US03-35733-10	Sequence 10, Appl
	4	3761	99.7	735	23	US-09-848-909A-10	Sequence 10, Appl
	5	3758	99.6	735	1	PCT-US03-19786-4	Sequence 4, Appl
	6	3755	99.5	735	1	PCT-US03-35733-2	Sequence 2, Appl
	7	3755	99.5	735	23	US-09-848-909A-2	Sequence 2, Appl
	8	3754	99.5	735	1	PCT-US03-35733-8	Sequence 8, Appl
	9	3754	99.5	735	23	US-09-848-909A-8	Sequence 8, Appl
	10	3753	99.4	735	1	PCT-US03-35733-13	Sequence 13, Appl
	11	3753	99.4	735	23	US-09-848-909A-13	Sequence 13, Appl
	12	3750	99.4	735	1	PCT-US03-35733-7	Sequence 7, Appl
	13	3750	99.4	735	23	US-09-848-909A-7	Sequence 7, Appl
	14	3749	99.3	735	1	PCT-US03-35733-4	Sequence 4, Appl
	15	3749	99.3	735	23	US-09-848-909A-4	Sequence 4, Appl
	16	3749	99.3	735	23	US-09-848-909A-4	Sequence 4, Appl
	17	3749	99.3	735	23	US-09-848-909A-6	Sequence 6, Appl
	18	3748	99.3	735	1	PCT-US03-35733-5	Sequence 5, Appl
	19	3748	99.3	735	1	PCT-US03-35733-19	Sequence 19, Appl
	20	3748	99.3	735	1	PCT-US03-35733-20	Sequence 20, Appl
	21	3748	99.3	735	1	PCT-US03-35733-21	Sequence 21, Appl
	22	3748	99.3	735	22	US-09-791-537-43735	Sequence 43735, A
	23	3748	99.3	735	23	US-09-848-909A-5	Sequence 5, Appl
	24	3748	99.3	735	23	US-09-848-909A-19	Sequence 19, Appl
	25	3748	99.3	735	23	US-09-848-909A-20	Sequence 20, Appl
	26	3748	99.3	735	23	US-09-848-909A-21	Sequence 21, Appl
	27	3748	99.3	735	30	US-10-410-647-30	Sequence 30, Appl
	28	3748	99.3	736	1	PCT-US01-14372A-1	Sequence 1, Appl
	29	3748	99.3	736	1	PCT-US01-14372A-2	Sequence 2, Appl
	30	3748	99.3	736	1	PCT-US01-14372A-3	Sequence 3, Appl
	31	3748	99.3	736	1	PCT-US01-14372A-4	Sequence 4, Appl
	32	3748	99.3	736	1	PCT-US01-14372A-5	Sequence 5, Appl
	33	3748	99.3	736	1	PCT-US01-14372A-6	Sequence 6, Appl
	34	3748	99.3	736	1	PCT-US01-14372A-7	Sequence 7, Appl
	35	3748	99.3	736	1	PCT-US01-14372A-8	Sequence 8, Appl
	36	3748	99.3	736	1	PCT-US01-14372A-9	Sequence 9, Appl
	37	3748	99.3	736	1	PCT-US01-14372A-10	Sequence 10, Appl
	38	3748	99.3	736	1	PCT-US01-14372A-11	Sequence 11, Appl
	39	3748	99.3	736	1	PCT-US01-14372A-12	Sequence 12, Appl
	40	3748	99.3	736	1	PCT-US01-14372A-13	Sequence 13, Appl
	41	3748	99.3	736	1	PCT-US01-14372A-14	Sequence 14, Appl
	42	3748	99.3	736	1	PCT-US01-14372A-15	Sequence 15, Appl
	43	3748	99.3	736	1	PCT-US01-14372A-16	Sequence 16, Appl
	44	3748	99.3	736	1	PCT-US01-14372A-17	Sequence 17, Appl
	45	3748	99.3	736	1	PCT-US01-14372A-18	Sequence 18, Appl

## ALIGNMENTS

## RESULT 1

PCT-US03-35733-11  
; Sequence 11, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; FILE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-11

Query Match 100.0%; Score 3774; DB 1; Length 735;  
Best Local Similarity 100.0%; Pred. No. 2.2e-298;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Db 1 EVQENRLLNESSSQGLLYFSDLPNFPQAPMVVTSSTTGDLSIPSSSELENIPESENYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDDRDN 180
Qy 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Db 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIDADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIKAKENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNIAATYNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNIAATYNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPENGL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPENGL 540
Qy 541 QYQKDIETEPDFNPDQSTQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Db 541 QYQKDIETEPDFNPDQSTQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

## RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication No. US2002003958A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; OF TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR APPLICATION DATE: 2001-05-04
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-14
```

```
Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 1 EVQENRLLNESSSQGLLYFSDLPNFPQAPMVVTSSTTGDLSIPSSSELENIPESENYF 60
Db 1 EVQENRLLNESSSQGLLYFSDLPNFPQAPMVVTSSTTGDLSIPSSSELENIPESENYF 60
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQPELKQKSSNSRKRSTASGTPVDDRDN 180
Qy 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Db 181 DGPIDSLVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSPKWKSTASDPYSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIDADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYNNVPTTSLVLGKNOTLATIKAKENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNIAATYNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKFSSTPIITMNNYQFLEKTKQLRLDTPQVYGNIAATYNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPENGL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNPENGL 540
Qy 541 QYQKDIETEPDFNPDQSTQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Db 541 QYQKDIETEPDFNPDQSTQNIKNQLAELNATNIYTVLDKIKLNAKWNILIRDKRPHYDR 600
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660
Qy 661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

Search completed: May 3, 2004, 20:02:11  
Job time : 38.6393 secs

Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTADPVSDEKVT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Qy 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQVYGNIAIYFNENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQVYGNIAIYFNENGRVVRVDTGSNWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Qy 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Qy 601 NNAVGADESVEKAEHREVNINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNAVGADESVEKAEHREVNINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKNTIINPSENGDTSTNG 720  
Qy 721 IKKILPSKKGYEIG 735  
Db 721 IKKILPSKKGYEIG 735

RESULT 13  
US-09-848-909-12  
; Sequence 12, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-12  
Query Match 99.3%; Score 3748; DB 12; Length 736;  
Best Local Similarity 99.5%; Pred. No. 5.1e-273;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
Qy 61 QSAIWSGFIKVKSDYEYFATSADNHYTMWDDQEVINKASNNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDYEYFATSADNHYTMWDDQEVINKASNNKIRLEKGRLYQIKIY 120

Qy 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKGSTASGPTVPDRN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKGSTASGPTVPDRN 180  
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTADPVSDEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTADPVSDEKVT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300  
Qy 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQVYGNIAIYFNENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKFFSTPTIMYNOFLEKTKQLRLDQVYGNIAIYFNENGRVVRVDTGSNWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Qy 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
Qy 601 NNAVGADESVEKAEHREVNINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNAVGADESVEKAEHREVNINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNYAVTKNTIINPSENGDTSTNG 720  
Qy 721 IKKILPSKKGYEIG 735  
Db 721 IKKILPSKKGYEIG 735

RESULT 14  
US-09-848-909-13  
; Sequence 13, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-13  
Query Match 99.3%; Score 3748; DB 12; Length 736;  
Best Local Similarity 99.5%; Pred. No. 5.1e-273;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60

Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKFSSTPTIMVYNQFLELEKTKQLRDLDDQVYGNATYVNGRVRVDTGSNWEV 480  
Db 421 LNAQDFSTPTIMVYNQFLELEKTKQLRDLDDQVYGNATYVNGRVRVDTGSNWEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Qy 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYVLDKIKLAKNNILIRDKRPHYDR 600  
Db 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYVLDKIKLAKNNILIRDKRPHYDR 600  
Qy 601 NNIAGADESVVKEAHREVINSGTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSGTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660  
Qy 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735  
RESULT 11  
US-09-848-909-10  
; Sequence 10, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-10  
Query Match 99.3%; Score 3748; DB 12; Length 736;  
Best Local Similarity 99.5%; Pred. No. 5.1e-273;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60  
Qy 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120  
Qy 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLPELKQKSNRKRSTAGTVPDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLPELKQKSNRKRSTAGTVPDRDN 180  
Qy 181 DGPISLVEGYTVDVQNKRTFLSPWISNIEHKKGLTKYKSPKWSFASDPYDFEKT 240  
Db 181 DGPISLVEGYTVDVQNKRTFLSPWISNIEHKKGLTKYKSPKWSFASDPYDFEKT 240  
Qy 241 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPKARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300

Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKFSSTPTIMVYNQFLELEKTKQLRDLDDQVYGNATYVNGRVRVDTGSNWEV 480  
Db 421 LNAQDFSTPTIMVYNQFLELEKTKQLRDLDDQVYGNATYVNGRVRVDTGSNWEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGFNEPKNL 540  
Qy 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYVLDKIKLAKNNILIRDKRPHYDR 600  
Db 541 QYQKDIITFDFNFDOQTSONIKQLAELNATNIYVLDKIKLAKNNILIRDKRPHYDR 600  
Qy 601 NNIAGADESVVKEAHREVINSGTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSGTEGLLNIDKIRKILSGYIIEDETEGLKEVINDRY 660  
Qy 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735  
RESULT 12  
US-09-848-909-11  
; Sequence 11, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-11  
Query Match 99.3%; Score 3748; DB 12; Length 736;  
Best Local Similarity 99.5%; Pred. No. 5.1e-273;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENOYF 60  
Qy 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120  
Qy 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLPELKQKSNRKRSTAGTVPDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLPELKQKSNRKRSTAGTVPDRDN 180  
Qy 181 DGPISLVEGYTVDVQNKRTFLSPWISNIEHKKGLTKYKSPKWSFASDPYDFEKT 240  
Db 181 DGPISLVEGYTVDVQNKRTFLSPWISNIEHKKGLTKYKSPKWSFASDPYDFEKT 240



```
Db 541 QYQCKDITFDNFQOQTSQNKQLAELNATNIYTVLDKIKNAKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNLSLRQDQKTFIDFKKYNKPLVYISNPNYKVNYYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNLSLRQDQKTFIDFKKYNKPLVYISNPNYKVNYYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 9
US-09-848-909-8
; Sequence 8, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-8

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKCRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKCRSTSGPTVPDRDN 180
QY 181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
QY 361 NANIRYNTGTAPYINVLPTTSLVLGKQNTLATIDADENQLSOLAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPYINVLPTTSLVLGKQNTLATIDADENQLSOLAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPIWNNYQFLEKTKQLRLDTPQVYGNIAATYFNGRVRVDTGSNWSEV 480
Db 421 LNAQKFSSTPIWNNYQFLEKTKQLRLDTPQVYGNIAATYFNGRVRVDTGSNWSEV 480
```

```
QY 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSPDLETTKPDMTLKEALKIAFGNEPNGML 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAANVPSPDLETTKPDMTLKEALKIAFGNEPNGML 540
QY 541 QYQCKDITFDNFQOQTSQNKQLAELNATNIYTVLDKIKNAKNILIRDKRPHYDR 600
Db 541 QYQCKDITFDNFQOQTSQNKQLAELNATNIYTVLDKIKNAKNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSGTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNLSLRQDQKTFIDFKKYNKPLVYISNPNYKVNYYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNLSLRQDQKTFIDFKKYNKPLVYISNPNYKVNYYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 10
US-09-848-909-9
; Sequence 9, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-9

Query Match 99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLLYYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKCRSTSGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKCRSTSGPTVPDRDN 180
QY 181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
Db 181 DGPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
QY 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
QY 361 NANIRYNTGTAPYINVLPTTSLVLGKQNTLATIDADENQLSOLAPNNYPSKNLAPIA 420
```

```
Db 721 IKKILIFSKGYEIG 735
|||||
RESULT 7
US-09-848-909-6
; Sequence 6, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-6

Query Match          99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWGFIKVKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWGFIKVKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVLSNDLQPELKQKSSNRKRSSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVLSNDLQPELKQKSSNRKRSSTAGTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTDPSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTDPSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQIILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFFSSTPTIMYVNFQLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKFFSSTPTIMYVNFQLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPKNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPKNL 540
Qy 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHDR 600
Qy 601 NNIAGADESVKAEHREVSINGTEGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVKAEHREVSINGTEGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660

US-09-848-909-7
; Sequence 7, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-7

Query Match          99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWGFIKVKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWGFIKVKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKEVLSNDLQPELKQKSSNRKRSSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKEVLSNDLQPELKQKSSNRKRSSTAGTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTDPSDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASTDPSDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQIILAPNNYPSKNLAPIA 420
Qy 421 LNAQKFFSSTPTIMYVNFQLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Db 421 LNAQKFFSSTPTIMYVNFQLELEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPKNL 540
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPKNL 540
Qy 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNKAMNIIIRDKEFHDR 600
Qy 601 NNIAGADESVKAEHREVSINGTEGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVKAEHREVSINGTEGLLNIDKIDKILSGYIVIEDETEGLKEVINDRY 660
```

; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-4

Query Match 99.3%; Score 3748; DB 12; Length 736;  
Best Local Similarity 99.5%; Pred. No. 5.1e-273;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSESSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSESSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDSYTATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDSYTATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180  
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240  
DB 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480  
DB 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAPFNEPNGNL 540  
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAPFNEPNGNL 540  
QY 541 QYQKDIETFDNFDDQSTQNIKNQLAELNATNIYVLDKIKLNKQNLILIRDKRPFHYDR 600  
DB 541 QYQKDIETFDNFDDQSTQNIKNQLAELNATNIYVLDKIKLNKQNLILIRDKRPFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 6  
US-09-848-909-5

; Sequence 5, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-5  
Query Match 99.3%; Score 3748; DB 12; Length 736;  
Best Local Similarity 99.5%; Pred. No. 5.1e-273;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSESSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSESSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDSYTATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDSYTATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLPQAPMVVTSSTGDLSPSSSELENIPSENQYF 180  
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240  
DB 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSASDPYSDPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480  
DB 421 LNAQKFSSTPTIMYNNQFLEKTKQLRLDQVYGNATYVNFENGRVVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAPFNEPNGNL 540  
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMLTKEALKIAPFNEPNGNL 540  
QY 541 QYQKDIETFDNFDDQSTQNIKNQLAELNATNIYVLDKIKLNKQNLILIRDKRPFHYDR 600  
DB 541 QYQKDIETFDNFDDQSTQNIKNQLAELNATNIYVLDKIKLNKQNLILIRDKRPFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735

US-09-848-909-2

```
Query Match          99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELQKSSNSRKRSTASGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELQKSSNSRKRSTASGPTVPDRDN 180
QY 181 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 181 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
QY 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVDTSNWSSEV 480
Db 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNITSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNITSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

RESULT 4

```
US-09-848-909-3
; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
```

```
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PPT
; ORGANISM: Bacillus anthracis
US-09-848-909-3
```

```
Query Match          99.3%; Score 3748; DB 12; Length 736;
Best Local Similarity 99.5%; Pred. No. 5.1e-273;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIQY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELQKSSNSRKRSTASGPTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELQKSSNSRKRSTASGPTVPDRDN 180
QY 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKT 240
Db 181 DGIPLDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASDPYSDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVDTSNWSSEV 480
Db 421 LNAQKFSSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVDTSNWSSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDQQTSONIKQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNITSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNITSSLRQDGKTFIDFKKYNKDLPLYISNPYKNVYAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735
```

RESULT 5

```
US-09-848-909-4
; Sequence 4, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
```

Db 61 QSAIWGFIKVKKSDYTTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASGTPVDDRDN 180  
Db 121 QRENPTKEGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASGTPVDDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTONTSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTONTSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420  
QY 421 LNAQKFSSTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
Db 421 LNAQKFSSTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
QY 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600  
Db 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600  
QY 601 NNTAVGADSVVKEAREHREINVESTGLLNIDKIRKILSGYIVEIETDGLKEVINDRY 660  
Db 601 NNTAVGADSVVKEAREHREINVESTGLLNIDKIRKILSGYIVEIETDGLKEVINDRY 660  
QY 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 720

## RESULT 2

US-09-848-909-1  
; Sequence 1, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 736  
; TYPE: PRN  
; ORGANISM: Bacillus anthracis  
US-09-848-909-1

Query Match 98.3%; Score 3748; DB 12; Length 736;  
Best Local Similarity 99.5%; Pred. No. 5.1e-273;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVQKQENLLNESSSQGLLYGYFSDLNFPQAPVVTSSITGDLISIPSSSELENIPSENOYF 60  
Db 1 EVQKQENLLNESSSQGLLYGYFSDLNFPQAPVVTSSITGDLISIPSSSELENIPSENOYF 60  
QY 61 QSAIWGFIKVKKSDYTTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120  
Db 61 QSAIWGFIKVKKSDYTTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIQY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASGTPVDDRDN 180  
Db 121 QRENPTKEGLDFKLYWTDSONKEVSSDNLOLPELKOKSSNSRKRSTASGTPVDDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTONTSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTONTSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAGFFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSOILAPNNYPSKNLAPIA 420  
QY 421 LNAQKFSSTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
Db 421 LNAQKFSSTPTIMYNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
QY 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600  
Db 541 QYQKDIETDFDFNDQDQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRFHYDR 600  
QY 601 NNTAVGADSVVKEAREHREINVESTGLLNIDKIRKILSGYIVEIETDGLKEVINDRY 660  
Db 601 NNTAVGADSVVKEAREHREINVESTGLLNIDKIRKILSGYIVEIETDGLKEVINDRY 660  
QY 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKNNVYAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYBIG 735  
Db 721 IKKILIFSCKGYBIG 735

## RESULT 3

US-09-848-909-2  
; Sequence 2, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 736  
; TYPE: PRN  
; ORGANISM: Bacillus anthracis

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 36.6393 Seconds  
(without alignments)  
5560.545 Million cell updates/sec

Title: US-09-848-909a-11

Perfect score: 3774

Sequence: 1 EVKQENLLNESSESSQGLL.....TSTNGIKKILFFSKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/prodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/prodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/prodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/prodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/prodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/prodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/prodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/prodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/prodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/prodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/prodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3748	99.3	735	15	US-10-410-647-30
2	3748	99.3	736	12	US-09-848-909-1
3	3748	99.3	736	12	US-09-848-909-2
4	3748	99.3	736	12	US-09-848-909-3
5	3748	99.3	736	12	US-09-848-909-4
6	3748	99.3	736	12	US-09-848-909-5
7	3748	99.3	736	12	US-09-848-909-6
8	3748	99.3	736	12	US-09-848-909-7
9	3748	99.3	736	12	US-09-848-909-8
10	3748	99.3	736	12	US-09-848-909-9
11	3748	99.3	736	12	US-09-848-909-10
12	3748	99.3	736	12	US-09-848-909-11
13	3748	99.3	736	12	US-09-848-909-12
14	3748	99.3	736	12	US-09-848-909-13
15	3748	99.3	736	12	US-09-848-909-14

16	3748	99.3	736	12	US-09-848-909-15	Sequence 15, Appl
17	3748	99.3	736	12	US-09-848-909-16	Sequence 16, Appl
18	3748	99.3	736	12	US-09-848-909-17	Sequence 17, Appl
19	3748	99.3	736	12	US-09-848-909-18	Sequence 18, Appl
20	3748	99.3	736	12	US-09-848-909-19	Sequence 19, Appl
21	3748	99.3	736	12	US-09-848-909-20	Sequence 20, Appl
22	3748	99.3	736	12	US-09-848-909-21	Sequence 21, Appl
23	3748	99.3	736	12	US-09-848-909-22	Sequence 22, Appl
24	3748	99.3	736	15	US-10-442-502-7	Sequence 7, Appl
25	3748	99.3	763	15	US-10-442-502-5	Sequence 5, Appl
26	3748	99.3	764	15	US-10-442-502-6	Sequence 6, Appl
27	3745	99.2	735	12	US-10-402-466A-9	Sequence 9, Appl
28	3745	99.2	735	12	US-10-402-466A-13	Sequence 13, Appl
29	3745	99.2	735	14	US-10-332-282-13	Sequence 13, Appl
30	3741	99.1	735	12	US-09-848-909-30	Sequence 30, Appl
31	3741	99.1	764	12	US-10-253-286-681	Sequence 681, App
32	3741	99.1	764	15	US-10-245-871-681	Sequence 681, App
33	3641	96.5	764	9	US-09-747-521-4	Sequence 4, Appl
34	3641	96.5	764	13	US-10-106-014-4	Sequence 4, Appl
35	3641	96.5	764	13	US-10-106-695-4	Sequence 4, Appl
36	3641	96.5	764	14	US-10-105-694-4	Sequence 4, Appl
37	3037	80.5	599	12	US-09-848-909-24	Sequence 24, Appl
38	3032	80.3	595	14	US-10-332-282-11	Sequence 11, Appl
39	2905	77.0	573	12	US-10-402-466A-22	Sequence 22, Appl
40	2893	76.7	569	15	US-10-442-502-8	Sequence 8, Appl
41	2483	65.8	487	14	US-10-332-282-5	Sequence 5, Appl
42	2168	57.4	426	14	US-10-332-282-9	Sequence 9, Appl
43	2122	56.2	423	12	US-10-402-466A-24	Sequence 24, Appl
44	1619	42.9	318	14	US-10-332-282-7	Sequence 7, Appl
45	1338	35.5	258	14	US-10-332-282-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-410-647-30  
; Sequence 30, Application US/10410647  
; Publication No. US20030235818A1  
; GENERAL INFORMATION:  
; APPLICANT: PLEXUS VACCINE, INC.  
; APPLICANT: Katritch, Vaevelod  
; APPLICANT: Bordner, Andrew  
; APPLICANT: Deans, Robert  
; APPLICANT: Sumner, Mary  
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME  
; FILE REFERENCE: PLEX1110-1  
; CURRENT APPLICATION NUMBER: US/10/410,647  
; CURRENT FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: US 60/373,668  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: US 60/371,256  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: US 60/371,250  
; PRIOR FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-10-410-647-30

Query Match	99.3%	Score 3748	DB 15	Length 735
Best Local Similarity	99.5%	Pred. No. 5.1e-273		
Matches 731	Conservative 0	Mismatches 4	Indels 0	Gaps 0
Qy	1	EVKQENLLNESSESSQGLLYFSDINFOAPMVVTSSTTGDLSIPSELENIPSENOYF	60	
Db	1	EVKQENLLNESSESSQGLLYFSDINFOAPMVVTSSTTGDLSIPSELENIPSENOYF	60	
Qy	61	QSAWTSFGFKVKSDYTFATSDNHNVTMWVDDQEVINKASNGNKIRLEGRLYQIKIQY	120	

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 884 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-471-044-5

Query Match 20.3%; Score 766; DB 2; Length 884;

Best Local Similarity 30.1%; Pred. No. 1.5e-46;

Matches 245; Conservative 129; Mismatches 295; Indels 146; Gaps 34;

```

QY      4 QENRLNSESQGLGYFSDNLFQAPMVVTSITGDLSPSELEN--IPSENYFFQ 61
Db      42 QXNQ---QXEMRKGLGYFKGDF-SNLTFAPTRDSTLIYDQOTANKLLDKQGEYQ 97

QY      62 SAIWSGFIKVKSDRYTFATSNHNHVTMMVDDQEVINKASNSNKIRLEKGRLYQIKYQ 121
Db      98 SIRWIGLIQSKETGDTFNLSEDEQAIIIEINGKIIISNKGKQVYVHLEKGLVPIKIEYQ 157

QY     122 RENPTEKGLD-----FKLYMTDSQNKKEVISSDNLQLPELKOKSS-----N 162
Db     158 SD--TKFNIDSTFKELKLFKIDSONQPOQVQODELRNPEFNKESQEFLEKPSKINLFT 215

QY     163 SRKKSSTAGTPVPRDNDGIPDSLEVEGYTVDVNKKTFPLSPWISNIHEKGLTKYKSS 222
Db     216 QXMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVKWDDSL-ASKGYTKFVSN 267

QY     223 PEKWSSTASPYSDFEKVTGRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDOSTQNT 282
Db     268 PLESHTVGDPYTDIEKARDLDSNAKETFNPLVAAFPSVNVSMKVLSPNENLS----- 323

QY     283 DSETRTISKNTSTRTHTSEVHGNAPVHASFFDIGGSVSAGFSNSNS---TVAIDHSL 339
Db     324 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHSET 366

QY     340 LAGERTWAEIMG---LNTADIPARNANRYVNTGTAPIYVLPITTSVLGKNOTLATID 395
Db     367 VAQE--WGSTGTGNTSFNTASAGYLNANRYNNVGTGALYDVKPTTSFVL-NDTIAIT 423

QY     396 ADENQLSQILAPNNYYPKSNLAPIALNAQKFSSTPTNNYNQFLELEKTKQLRLDTQV 455
Db     424 AKSNSTALNISPGEYPKKGQNGIAITSMDDFNHPITLNNKQVDNLLNNKPMLETNQT 483

QY     456 YGNIAITYNFENGRVVDTSNWEVLPOIQTETARIIFNGKDLNLVERRIAIVNPSDPLE 515
Db     484 DG---VYKIKDTRHGNVITGEMNGVIOIKAKTASIIIVDDGE--RVAERKVAARKYENPED 539

QY     516 TTKPDMTLKALKIAF--GFNEPENGVLQCGKDITEFDF--NPDQOTSQNIKNQLAEL-- 569
Db     540 KT-PSLTKDALKLSYDPEIKEIEGLLYYKPKFIESSVMTYLDENTAKEVTKQLNDITG 598

QY     570 ---NATNIYTVLDKIKLNARONILIRDKRFHYDRNNIANGADESVVKEAHRVINSSTEG 626
Db     599 KFKDVSHLYDV---KLTPOANTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGNG 650

QY     627 -----LLNID-----KDIRKILSGYIIVEIDTE-----GLKE 654
Db     651 KQYSSNNPANTLNTDAQELKNKRDYVYISLYMKSEKNTQCEITIDGEIYPIITKTVN 710

QY     655 VINDRYDMLN-----ISSLRQGGTFIDFKYNDKPLIYISNPNYKVNVAATKE 704
Db     711 VNKNYKRLDIIAHNKSNPISLH-----IKTNBETILFWDDDISI-TDVASIKPE 760

QY     705 NTIINPSENGDT-STNGIK---KILLFSGKGYEIG 735
Db     761 N--LTDSEIKQIYRYGKLEKEDGILIDKGGIHYG 793

```

Search completed: May 3, 2004, 19:42:45

Job time : 18.0986 secs

516 TTKPMTLKEAKIAF--GFNEPNGNLYOQOKDITEFDF--NFDOQTSONIKKOLAEL-- 569  
540 KT-PSLTLDKALSLYDEIEXIEGLLYYKVKPIYESVWYLDENTAKETVKQNDITG 598  
570 ---NATNIYTVLDKIKNAKNILIRDKRPHYDRNNIAGADESVVKEAHEVINSSTEG 626  
599 KFKVSHLYDV---KLTFFKNVVIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 650  
627 -----LLNID-----KDRKILSGYIVEIEDTE-----GLKE 654  
651 KQYSSNNPDANLTLNTDAQEKLNKNRDY--SLYMKSEKNTCEITIDGEIYPTTKTVN 710  
655 VINDRYDMLN-----TSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAATKE 704  
711 VNKNYKRLDIIAHNIKSNPSSLH-----IKTNDIITLFWDDISI-TDVASIKPE 760  
705 NTIINPSENGDT-STNGIK---KILIFSKKGYEIG 735  
761 N-LTDSKIKYISRYGKLEKGLIDKKGHIYG 793

RESULT 15  
US-08-471-044-5  
Sequence 5, Application US/08471044  
Patent No. 5840868  
GENERAL INFORMATION:  
APPLICANT: Warren, Gregory W  
APPLICANT: Koziel, Michael G  
APPLICANT: Mullins, Martha A  
APPLICANT: Nye, Gordon J  
APPLICANT: Carr, Brian  
APPLICANT: Desai, Nalini M  
APPLICANT: Kostichka, N. Kristy  
APPLICANT: Duck, Nicholas B  
APPLICANT: Estruch, Juan J  
TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESS: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,044  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/463,483  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/314,594  
FILING DATE: 09-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/218,018  
FILING DATE: 23-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/037,057  
FILING DATE: 25-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Gary M.  
REGISTRATION NUMBER: 40,403  
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLV3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8582  
TELEFAX: 919-541-8689

516 TTKPMTLKEAKIAF--GFNEPNGNLYOQOKDITEFDF--NFDOQTSONIKKOLAEL-- 569  
540 KT-PSLTLDKALSLYDEIEXIEGLLYYKVKPIYESVWYLDENTAKETVKQNDITG 598  
570 ---NATNIYTVLDKIKNAKNILIRDKRPHYDRNNIAGADESVVKEAHEVINSSTEG 626  
599 KFKVSHLYDV---KLTFFKNVVIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 650  
627 -----LLNID-----KDRKILSGYIVEIEDTE-----GLKE 654  
651 KQYSSNNPDANLTLNTDAQEKLNKNRDY--SLYMKSEKNTCEITIDGEIYPTTKTVN 710  
655 VINDRYDMLN-----TSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAATKE 704  
711 VNKNYKRLDIIAHNIKSNPSSLH-----IKTNDIITLFWDDISI-TDVASIKPE 760  
705 NTIINPSENGDT-STNGIK---KILIFSKKGYEIG 735  
761 N-LTDSKIKYISRYGKLEKGLIDKKGHIYG 793

Query Match 20.3%; Score 766; DB 1; Length 884;  
Best Local Similarity 30.1%; Pred. NO. 1.5e-46;  
Matches 245; Conservative 129; Mismatches 295; Indels 146; Gaps 34;

4 QENRLNSESQGLYVSDNFQAPMVVTSITGDLSPSELEN--IPSENQVFO 61  
42 QKQK---QKMDKGLLYFGKDF-SNLTFAPTRSTLIYDQOTANKLLDKKQBEYQ 97  
62 SAIWSGFIKVKSDYTFATSADNHVTMVVDQEVINKASNKIRLEKGLYQIKIYQ 121  
98 SIRWGLIQSKETGDTFNLSEDEQAIIBINGKIISNKGKQVHVHLEKGLVPIKIEYQ 157  
122 RENPTEKGLD-----FKLYWTDSONKKEVISDNLQPELKQKSS-----N 162  
158 SD--TKFNIDSKTFKELKLFKIDSNQPOQVQODELRNPEFNKESQEFLLAKPSKINLFT 215  
163 SRKGRSTAGPTVPDRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNTHKGLTKYKSS 222  
216 QMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 267  
223 PEKWTASDPYSPFKVGRIDKNVSPARHPLVAAPIVHVDMENILSKNEDOSTQNT 282  
268 PLESHTVGDPTDYEXAARDLDSNAKETFNPLVAAPSVNVSMKVLSPNENUS--- 323  
283 DSETRISKNTSTSRHTSEVHGNAHVHASFIDIGSVSAGFSNNGSS---TVADHSL 339  
324 ----NVESHSSNTWYNT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 366  
340 LAGERTWAETWG----LNTADTARLANRYNTGAPINVLPTTSVLVGNQIATID 395  
367 VAQE--WGTSTGNTSOFNTASAGYLNANRYNVGTGAIYDVKPTTSFVL--NNDTIAT 423  
396 ADENQLSQIILANNYPYKSLAPIALNAOKKFSPTITWNYNQFLELEKTKQLRLDQV 455  
424 AKNSIALNISPESYPKQNGIATSDMDFNSHPTITLNKKQVDNLNKNKPMWLENTQ 483  
456 YGNIATYFNENGRVVDTSNWSSEVLFPQIQTETARIIFNGKOLNVLVERIAAVNPDP 515  
484 DG---VYKIDKTHGNIVTGGEWNGVIOQIKAKTASIIIVDDGE-RVAEKRAAKDENP 539



Db	764	IKQVSRYGKLEBGLIDKGGIHYG	790
Db	95	SIRWIGLIQSKETGDTFTNLSEDEQAIIIEINGKILSNKGEKQVHVHLEKGLVPIKIEYQ	150
QY	122	RENPTKEGLD-----FKLYWTDSONKEVLISSNQLQPELKQSS-----N	162
Db	155	SD--TKFNIDSKTFKELKLFKIDSONQPOQVQDDELAPNPFNKESQBFLLAKPSKINLFT	212
QY	163	SRKKESTAGTPVPRDNDGIPDSLEVBEGYTDVKNKRTPLSPWISNHEKKGJTKYKSS	222
Db	213	QMKREIDED---TDTGDSIPDLWEENGYYI-----QNRIVAKVDDSL-ASKGYTKVSVN	264
QY	223	PEKWTASDPYDFEKFVTGRIDKNVSPPEARHPLVAAPVPIVHVDVENILSKNEDQSTQNT	282
Db	265	PLESHTVGDPYTDYEKAARDLDSNAKETFNPLVAAPFVNVSVNEKVLSPNENLS----	320
QY	283	DSETRTISKVTSRTHSEVHGNAEVHASFPDGGVSGVSAFSGSNSNS--TVAIDHSL	339
Db	321	-----NSVESHSSTWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS	363
QY	340	LAGEWTAEWTG---LNTADTARLNANIRVNTGTAPVYVNVLTPTTSLVLGKNOTLATID	395
Db	364	VAQE--WGTSTGNTSQFNTASAGYLNANVNYVNTGTGAIYDVKPTTSPVL--NNDTIAIT	420
QY	396	ADENQLSQILAPNNYPPSKNLAPALNAOKKFSSTPTMTNYNQFLELEKTKQLRLDITDQV	455
Db	421	AKSNSTALNIPGESYPKKGQNGIAITSMDDFNHSPITLKKQVDNLLNNKPMLEINQ	480
QY	456	YGNATYFENGVRVDTGSKWSEVLPOIQETTARIIFNGKDLNLVERRIAAAPSDDPLE	515
Db	481	DG---VYKIKDTHGNIVTGGEMNGVIOQIKAKTASIIYDDGE--RVAEKRVAAKDYENPED	536
QY	516	TTKPDMLTKEALKIAP--GFNEPENGNOYQOGKIDTEFDF--NFQOQTSQNTKNQLAEL--	569
Db	537	KT-PSLTLKDALKLSYDPDEIKIEGLLYYKKNKPIVSSWMTYLDENTAKEVTKQLNDITG	595
QY	570	NATNYTVLDKIKLNAKONILIRDFHYDRNNAVGADESUVVKEAHEVINSSTEG	626
Db	596	KPKDVSHLYDV---KLTPKMNVTIK-LSIILYDN---AESNDNSIGKWTNTNIVSGGKNG	647
QY	627	-----LLNID-----KDIRKILSGYIVEIETDE-----GLKE	654
Db	648	KQYSSNNPDANLTLNTDAQEKLNKRDYIISLYMKSEKTCETIDGIEYPIITKTVN	707
QY	655	VINDRYDMLNI--SSLRQDGKTFIDFKYNDKLPYISNPNYKNVYAVTNTIINSE	712
Db	708	VKNQNYKRLDIIAHNIKGNPISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDS	763
QY	713	NGDT-STNGIK---KILIFSKGYBIG	735
Db	764	IKQVSRYGKLEBGLIDKGGIHYG	790
RESULT 14			
US-08-471-033-5			
; Sequence 5, Application US/08471033			
; Patent No. 5770696			
; GENERAL INFORMATION:			
; APPLICANT: Warren, Gregory W			
; APPLICANT: Koziel, Michael G			
; APPLICANT: Mullins, Martha A			
; APPLICANT: Nye, Gordon J			
; APPLICANT: Carr, Brian			
; APPLICANT: Desai, Nalini M			
; APPLICANT: Kostichka, N. Kristy			
; APPLICANT: Duck, Nicholas B			
; APPLICANT: Estruch, Juan J			
; TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains			
; NUMBER OF SEQUENCES: 50			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: CIBA-GEIGY Corporation			
; STREET: 7 Skyline Drive			
; CITY: Hawthorne			
; STATE: NY			
; COUNTRY: USA			
US-09-850-351A-32			
Query Match 20.3%; Score 768; DB 4; Length 881;			
Best Local Similarity 30.1%; Pred. No. 1.1e-46;			
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;			
QY	4	QENLLNSESQGLGCVFSDNLFQPMVVTSTTGDLSIPSELEN--IPSENQYFO	61
Db	39	QKQ--QKMDRKGGLGIFKGD--SNLTFAPTRDSTLIYQCTANKLKKQGEYQ	94
QY	62	SAIWGCFIKVKSDBYTFATSNHVTWVDDQEVINKASNSKIRLEKGRYQIKYQ	121

QY 456 YGNATYNFENGRVVDTSNWSSEVLPOIQTETARIIFNGKDLNVERIAAANPSDPLE 515  
Db 481 DG---VYKIDKTHGNIVTGGWNGVIOOIKAFTASIIIVDDGE-RVAEKRAAKOYENPED 536  
QY 516 TTKPDMTLKEALKIAP--GFNEPENGNOYQOGKDIETBDF--NPDQOTSONIKNQLAEL-- 569  
Db 537 KT-PSLTLDKALKSYDPEIKETIBGLLYKKNKPIYESSVMTYLDENTAKEVTKQLNDITG 595  
QY 570 ---NATNIYTVLDKILNAKXNIIIRDRPHYDRNNAIAGADESVVKEAHREVINSSTEG 626  
Db 596 KFXDVSHLYDV---KLTPEMNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGG 647  
QY 627 -----LLLNID-----KOIRKILSGYIYEIETDE-----GLKE 654  
Db 648 KKOYSSNNPDANLTNTDAQKLNKRDYIISLYMKSEKTOCEITIDGIEIYPIITKTVN 707  
QY 655 VINDRYDMLNI---SSLQDQKTFIDPKYNDKPLIYSNPNYKVNYYAVTKENTINPSE 712  
Db 708 VNKDNYKRLDIIAHNIKSNPISIIHKT-NDBITLFWDDISI--TDVASIKPEN--LTDSE 763  
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735  
Db 764 IKQIYRYGKLEDDGILIDKKGIIHYG 790

RESULT 12  
US-09-307-106-8  
; Sequence 8, Application US/09307106  
; Patent No. 6603063  
; GENERAL INFORMATION:  
; APPLICANT: Feitelson, Jerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schweits, James  
; APPLICANT: Loewer, David  
; APPLICANT: Duilum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; APPLICANT: Morrill, George  
; APPLICANT: Finstad-Lee, Stacey  
; TITLE OF INVENTION: No. 6603063a1 Pesticidal Toxins and Nucleotide  
; TITLE OF INVENTION: Sequences Which Encode These Toxins  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32608-6669  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/307,106  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,848  
; FILING DATE: 30-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/960,780  
; FILING DATE: 30-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/073,898  
; FILING DATE: 05-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sanders, Jay M.  
; REGISTRATION NUMBER: 39,355

; REFERENCE/DOCKET NUMBER: MA-708C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 881 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: PS177C8a  
; US-09-307-106-8

Query Match 20.3%; Score 768; DB 4; Length 881;  
Best Local Similarity 30.1%; Pred. No. 1,1e-46;  
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;  
QY 4 QENRLNBSSESSQGLLYGYSFSLNFPQAPMVTSSTGDLSPSSSELEN--IPSENQVPO 61  
Db 39 QKIQ---QKEMDRKGLLYGFKGDF-SNLTFAPTRDSTLIYDQOTANKLDDKKQOEYQ 94  
QY 62 SAIWSGFIKVKSDDEVTFATSDNHNVTMVMDQEVINKASNNKIRLEKRLYQIKIOYQ 121  
Db 95 SIRNIGLIQSKETGDFTNLSEDEQAIIEINGKIISNKGKEQVHVLEKGLVPIKIHQ 154  
QY 122 RENPTKGLD-----FKLYWTDSONKKEVSSDNLQIPELKQKSS-----N 162  
Db 155 SD--TKFNIDSKTFKELKLFKIDSONQPOQOQDELNPEFNKESQBFKAPKSNLFT 212  
QY 163 SRKRSTSGAPTVDRDNDGIPDSLEVEGYTVVKNKRTFLSPWISNTHKGLTKYSS 222  
Db 213 QMKREIDED---TDTGDSIPDLNEENGYTI---QNRIVAKWDDSI-ASKGYTKFSN 264  
QY 223 PEKWTASDPYDFEKVTRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDOSTQNT 282  
Db 265 PLESHTVGDPTDYKAAARDLDSNAKETFNPLVAAPFSVNVSMKVLSPNENLS---- 320  
QY 283 DSETRITKNTSTSRTHTSEVHGNABVHASPFDIGSVSAGFSNSNS---TVAIDHSL 339  
Db 321 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGSVNYQHS 363  
QY 340 LAGERWAEITMG---LNTADTARLNANRYNTGTAPIYVLTPTSLVLGKQNTLATID 395  
Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANRYNVNVTGAIYDVKPTTSFVL--NNDTIA 420  
QY 396 ADENQLSQILAPNYYPSKNLAPIALNAOKFSSPTIMVYNOFLEKTKOLRLDTQV 455  
Db 421 AKSNSTALNISPGESYPKKGQNGIATSMDDDFNSHPITLNKKQVDNLLNKKPMLE 480  
QY 456 YGNATYNFENGRVVDTSNWSSEVLPOIQTETARIIFNGKDLNVERIAAANPSDPLE 515  
Db 481 DG---VYKIDKTHGNIVTGGWNGVIOOIKAFTASIIIVDDGE-RVAEKRAAKOYEN 536  
QY 516 TTKPDMTLKEALKIAP--GFNEPENGNOYQOGKDIETBDF--NPDQOTSONIKNQLAEL-- 569  
Db 537 KT-PSLTLDKALKSYDPEIKETIBGLLYKKNKPIYESSVMTYLDENTAKEVTKQLNDITG 595  
QY 570 ---NATNIYTVLDKILNAKXNIIIRDRPHYDRNNAIAGADESVVKEAHREVINSSTEG 626  
Db 596 KFXDVSHLYDV---KLTPEMNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGG 647  
QY 627 -----LLLNID-----KOIRKILSGYIYEIETDE-----GLKE 654  
Db 648 KKOYSSNNPDANLTNTDAQKLNKRDYIISLYMKSEKTOCEITIDGIEIYPIITKTVN 707  
QY 655 VINDRYDMLNI---SSLQDQKTFIDPKYNDKPLIYSNPNYKVNYYAVTKENTINPSE 712  
Db 708 VNKDNYKRLDIIAHNIKSNPISIIHKT-NDBITLFWDDISI--TDVASIKPEN--LTDSE 763  
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735

Db 95 SIRWIGLQSKETGDFNLSDEQAIIEINGKIIISNKGKQVHLEKGLVPIKIEYQ 154  
Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNTLQPLKOKSS-----N 162  
Db 155 SD--TKFNIDSKTFKELKFKIDSQOQVQOQDELNPFNFKESQEFLLAKPSKINFT 212  
Qy 163 SRKRSTASDPYSDPEKVTGTRDKNVSPARHPLVAAPYIVHVDMENIILSKNEDOSTQNT 282  
Db 213 QKMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264  
Qy 223 PEKMWSTASDPYSDPEKVTGTRDKNVSPARHPLVAAPYIVHVDMENIILSKNEDOSTQNT 282  
Db 265 PLESHTVGDPYTDYKAAARDLDSNAKETFPNLAAPFVSNVSKVILSPENLS---- 320  
Qy 283 DSETRTSKNTSTRTSTHTEVHGNAEVHASFDDIGSVSAGFSNSNS---TVAIDHSL 339  
Db 321 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSNVYQHS 363  
Qy 340 LAGERTWAETMG---LNTADTARLANIRVYNTGTAPYVNLPTTSLVLGKQTLATID 395  
Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANVRNNVGTGAIYDVKPTTSFVL--NNDTIAIT 420  
Qy 396 ADENQLSQILAPNNYPSKNAPIALNAOKKFSSTPTMKYNOFLELEKTKQLRLDTQV 455  
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNSHPITLKKQVDNLLNKPMLTNTQ 480  
Qy 456 YGNATATNFENGRVYDTGNSWSEVLPOIQTETARTIIFNGKDLNVLREIRAAVNSDPLE 515  
Db 481 DG---VYKIKTHGNITVGSENGVIOIIRAKTASIIVDGGE--RVAEKVAAKOVENPED 536  
Qy 516 TTKPDMTLKEALKIAP--GFNEPENGNIQYQKIDTEPDE--NPDQOTSONIKNQLAEL-- 569  
Db 537 KT-PSLTKDALKSYPEDEKETEGLLYKKNKPIYESSVMYLDENTAKEVTKQLNDITG 595  
Qy 570 ---NATNIYTVLDKIKNAXMILIRDKRPHYDNNIACADESVVKEAHEVINSSTEG 626  
Db 596 KFKOVSHLDYV---KUTPKMNVTK--LSILYN---AESNDNSIGKWTNIVSGGNG 647  
Qy 627 ---LNLNID-----KDIRKILGYVIEIDTE-----GLKE 654  
Db 648 KKOYSSNNPDANLTLNDAOKLNKRDYVYISLYMKSEKNTQCEITIDGIEIYPIITKT 707  
Qy 655 VINDRYDMLNT--SSLRQDGKTRIDFKYNDKPLVYISNPKYVNYVATKNTIINPSE 712  
Db 708 VKNYKXELDIANIKSNPISSHIKT--NDEITLFWDDISI--TDVASIKPEN--LTDSE 763  
Qy 713 NGDT-STNGIK---KILIFKKGVEIG 735  
Db 764 IKQIYRYGIKLEDGILIDKKGGIHYG 790

## RESULT 11

US-09-073-898-32

Sequence 32, Application US/09073898

Patent No. 6242669

GENERAL INFORMATION:

APPLICANT: Feitelson, Jerald S.

APPLICANT: Schneff, H. Ernest

APPLICANT: Narva, Kenneth E.

APPLICANT: Stockhoff, Brian A.

APPLICANT: Schmitts, James

APPLICANT: Loewer, David

APPLICANT: Dullum, Charles Joseph

APPLICANT: Muller-Cohn, Judy

APPLICANT: Stamp, Lisa

APPLICANT: Morrell, George

APPLICANT: Finstad-Lee, Stacey

TITLE OF INVENTION: No. 6242669el Pesticidal Toxins and Nucleotide

TITLE OF INVENTION: Sequences Which Encode These Toxins

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd &amp; Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/073,898  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,848  
FILING DATE: 30-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/960,780  
FILING DATE: 30-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-708C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 881 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: PS177C8  
US-09-073-898-32

Query Match 20.3%; Score 768; DB 3; Length 881;

Best Local Similarity 30.1%; Pred. No. 1.1e-46;

Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;

Qy 4 QENRLNSESQGLGYFSDLNFPQAPMVYTSSTGDLSPGSELEN--IPGENOVYQ 61  
Db 39 QKQK---QKEMRKGLLYGFFKQDF-SNLTFATRDSTLLYDOQTANKLLDKKQEQY 94  
Qy 62 SAIWGFIKVKSEDEYTFATSDNHNVTWDDQEVINKASNSNKIRLEKRLYQIKQYQ 121  
Db 95 SIRWIGLQSKETGDFNLSDEQAIIEINGKIIISNKGKQVHLEKGLVPIKIEYQ 154  
Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNTLQPLKOKSS-----N 162  
Db 155 SD--TKFNIDSKTFKELKFKIDSQOQVQOQDELNPFNFKESQEFLLAKPSKINFT 212  
Qy 163 SRKRSTASDPYSDPEKVTGTRDKNVSPARHPLVAAPYIVHVDMENIILSKNEDOSTQNT 282  
Db 213 QKMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264  
Qy 223 PEKMWSTASDPYSDPEKVTGTRDKNVSPARHPLVAAPYIVHVDMENIILSKNEDOSTQNT 282  
Db 265 PLESHTVGDPYTDYKAAARDLDSNAKETFPNLAAPFVSNVSKVILSPENLS---- 320  
Qy 283 DSETRTSKNTSTRTSTHTEVHGNAEVHASFDDIGSVSAGFSNSNS---TVAIDHSL 339  
Db 321 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSNVYQHS 363  
Qy 340 LAGERTWAETMG---LNTADTARLANIRVYNTGTAPYVNLPTTSLVLGKQTLATID 395  
Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANVRNNVGTGAIYDVKPTTSFVL--NNDTIAIT 420  
Qy 396 ADENQLSQILAPNNYPSKNAPIALNAOKKFSSTPTMKYNOFLELEKTKQLRLDTQV 455  
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNSHPITLKKQVDNLLNKPMLTNTQ 480

```
QY 237 EKVTGIDKNVSPARHPPLVAAPIVHVDMENIILSKNEDOSTONTSETRTISKNTSTS 296
Db 241 EKVTGIDKNVSPARHPPLVAAPIVHVDMENIILSKNEDOSTONTSETRTISKNTSTS 300
QY 297 RHTTSEVHGNAEVAHSAFFDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTAD 356
Db 301 RHTTSEVHGNAEVAHSAFFDGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTAD 360
QY 357 TARLNNIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENOLSQLAPNYYPSKNL 416
Db 361 TARLNNIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQSQLAPNYYPSKNL 420
QY 417 APIALNAQKFFSTPTTMYNNOFLEKTKQLRLDTQVYGNIAFYNFENGRVRYDTGSN 476
Db 421 APIALNAQDFFSTPTTMYN-----YGNIAFYNFENGRVRYDTGSN 460
QY 477 WSEVLPOIETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARFNEP 536
Db 461 WSEVLPOIETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARFNEP 520
QY 537 NGNLOQOGKDITEFDFNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNKXNILLIRDRKF 596
Db 521 NGNLOQOGKDITEFDFNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNKXNILLIRDRKF 580
QY 597 HYDRNNIAGDESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVI 656
Db 581 HYDRNNIAGDESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVI 640
QY 657 NDRYDMNLNSSLRQDGKTFIDFKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDT 716
Db 641 NDRYDMNLNSSLRQDGKTFIDFKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDT 700
QY 717 STNGIKKILIFSKKGYEIG 735
Db 701 STNGIKKILIFSKKGYEIG 719
RESULT 9
US-09-273-839A-8
; Sequence 8, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-273-839A-8
Query Match 33.8%; Score 1275; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 9.5e-84;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 486 ETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARFNEPNGNLQYQK 545
Db 29 ETTARIIFNGKDLNVERRIAANVPDPLETTKPDMTLKEALKIARFNEPNGNLQYQK 88
QY 546 DITEFDFNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNKXNILLIRDRGFHYDRNNIIV 605
Db 89 DITEFDFNFQOOTSQNIKNQLAELNATNIYTVLDKIKLNKXNILLIRDRGFHYDRNNIIV 148
QY 606 GABESVVKAEHREVINSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNI 665
Db 149 GABESVVKAEHREVINSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRYDMLNI 208
```

```
QY 666 SSIRQDGKTFIDPKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDTSTNGIKKIL 725
Db 209 SSIRQDGKTFIDPKYNDKPLVYISPNYKVVYAVTKENTIINPSENGDTSTNGIKKIL 268
QY 726 IFSKKGYEIG 735
Db 269 IFSKKGYEIG 278
RESULT 10
US-08-960-780-32
; Sequence 32, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 177C8
; US-08-960-780-32
Query Match 20.3%; Score 768; DB 3; Length 881;
Best Local Similarity 30.1%; Pred. No. 1.1e-46;
Matches 243; Conservative 134; Mismatches 300; Indels 130; Gaps 34;
QY 4 QENLLNESSSQGLGYYFSDLNFOAPMVVYVSTTGDLSIPSELEN--IPSENYFQ 61
Db 39 QKNO---QKMDRKGLGYYFKGKDF-SLNTMFAPTRDSTLIYDQQTANKLLDKKQEQY 94
QY 62 SAIWSGFIKVKSDSEYFATFADNHVTWVDDQEVINKASNSKIRLEKGLYQIKYQ 121
```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-31

Query Match 94.6%; Score 3571; DB 1; Length 719;  
Best Local Similarity 95.0%; Pred. No. 8.1e-248;  
Matches 702; Conservative 4; Mismatches 9; Indels 24; Gaps 2;  
QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDVEYFATSDADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDVEYFATSDADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVTSADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 176  
DB 121 QRENTEKGLDFKLYWTDSONKKEVTSADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 176  
QY 177 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNHEKGLTKYKSSPEKWTASDPYSDF 236  
DB 177 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNHEKGLTKYKSSPEKWTASDPYSDF 236  
QY 237 EKVTRIDKNVSPARPLVAAPFIVHVDMENILSKNEDQSTQNTDSETRISKNYTS 296  
DB 237 EKVTRIDKNVSPARPLVAAPFIVHVDMENILSKNEDQSTQNTDSETRISKNYTS 296  
QY 297 RHTSEVHGNAEVSFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGNTAD 356  
DB 297 RHTSEVHGNAEVSFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGNTAD 356  
QY 357 TARLNANRYNTGTAPYVNLPTSLVGLKNOTLATIDADENQLSOILAPNNYPSKNL 416  
DB 357 TARLNANRYNTGTAPYVNLPTSLVGLKNOTLATIDADENQLSOILAPNNYPSKNL 416  
QY 417 APIALNAQKFSSTPITMYNQFLELEKTKQLRLDTQVYGNATYNFENGRVRVDTGSN 476  
DB 417 APIALNAQKFSSTPITMYNQFLELEKTKQLRLDTQVYGNATYNFENGRVRVDTGSN 476  
QY 477 WSEVLPQIOETARIIIFNGKOLNVERIAAVNPSPDLETTKPDWTLKEALKIATGFNEP 536  
DB 477 WSEVLPQIOETARIIIFNGKOLNVERIAAVNPSPDLETTKPDWTLKEALKIATGFNEP 536  
QY 537 NGNLYQYQKDIETEDFNFDQTSQNIKNQLAELNATNIYVLDKIKLNANILIRDKRF 596  
DB 537 NGNLYQYQKDIETEDFNFDQTSQNIKNQLAELNATNIYVLDKIKLNANILIRDKRF 596  
QY 597 HYDRNNTAVGADESVEKAEHREVNSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVI 656  
DB 597 HYDRNNTAVGADESVEKAEHREVNSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVI 656

QY 657 NDYDMLNITSSLRQDGKTFIDFKKNDKULPLVSNPNYKVNKYAVTKENTINPSENGDT 716  
DB 641 NDYDMLNITSSLRQDGKTFIDFKKNDKULPLVSNPNYKVNKYAVTKENTINPSENGDT 700  
QY 717 STNGIKKILIFSKKGYEIG 735  
DB 701 STNGIKKILIFSKKGYEIG 719  
RESULT 8  
PCT-US94-01624-31  
Sequence 31, Application PC/TUS9401624  
GENERAL INFORMATION:  
APPLICANT: Leppla, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
TITLE OF INVENTION: RELATED METHODS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW  
STREET: Steuart Street Tower, 20th Floor, One Market  
STREET: Plaza  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/01624  
FILING DATE: June 25, 1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01624-31

Query Match 94.6%; Score 3571; DB 5; Length 719;  
Best Local Similarity 95.0%; Pred. No. 8.1e-248;  
Matches 702; Conservative 4; Mismatches 9; Indels 24; Gaps 2;  
QY 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
DB 1 EVKQENRLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSPSSSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDVEYFATSDADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDVEYFATSDADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVTSADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 176  
DB 121 QRENTEKGLDFKLYWTDSONKKEVTSADNHVMTWDDQEVINKASNKIRLEKGLYQIKIY 176  
QY 177 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNHEKGLTKYKSSPEKWTASDPYSDF 236  
DB 177 DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNHEKGLTKYKSSPEKWTASDPYSDF 236  
QY 237 EKVTRIDKNVSPARPLVAAPFIVHVDMENILSKNEDQSTQNTDSETRISKNYTS 296  
DB 237 EKVTRIDKNVSPARPLVAAPFIVHVDMENILSKNEDQSTQNTDSETRISKNYTS 296

Db 421 LNAQDDFSPTITMNVNQFLEKTKQLRLDQVYGNIAVYVNGRVRVDTGSMSEV 480  
 QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540  
 Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540  
 QY 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNKMMILIRDKRPHYDR 600  
 Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNKMMILIRDKRPHYDR 600  
 QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660  
 Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660  
 QY 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTKNTIINPSENGDTSTNG 720  
 Db 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTKNTIINPSENGDTSTNG 720  
 QY 721 IKKIL---IFSCKG 731  
 Db 721 IKKILKKVVLGKG 734

## RESULT 6

PCT-US94-01624-12  
 ; Sequence 12, Application PC/TUS9401624  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leppa, Stephen H.  
 ; APPLICANT: Klimpel, Kurt R.  
 ; APPLICANT: Arora, Naveen  
 ; APPLICANT: Singh, Yogendra  
 ; APPLICANT: Nichols, Peter J.  
 ; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
 ; TITLE OF INVENTION: RELATED METHODS  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW  
 ; STREET: Steuart Street Tower, 20th Floor, One Market  
 ; CITY: San Francisco  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94105  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US94/01624  
 ; FILING DATE: June 25, 1993  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Weber, Kenneth A.  
 ; REGISTRATION NUMBER: 31,677  
 ; REFERENCE/DOCKET NUMBER: 15280-115  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 543-9600  
 ; TELEFAX: (415) 543-5043  
 ; INFORMATION FOR SEQ ID NO: 12:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 903 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; PCT-US94-01624-12

Query Match 98.1%; Score 3703.5; DB 5; Length 903;  
 Best Local Similarity 98.6%; Pred. No. 3.5e-257;  
 Matches 724; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 EVKQENLLNESSSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
 |||||

Db 1 EVKQENLLNESSSQGLLYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
 QY 61 QSAIWGFTKVKKSDSYTTFATSDADNHTVWVDDQEVINKASNSNKIRLEKGLYQIKIY 120  
 Db 61 QSAIWGFTKVKKSDSYTTFATSDADNHTVWVDDQEVINKASNSNKIRLEKGLYQIKIY 120  
 QY 121 QRENPTKEGLDFKLYWTDSONKKEVSISSDNQLPELKQKSSNSRKKRSTSAOPTVDPDRN 180  
 Db 121 QRENPTKEGLDFKLYWTDSONKKEVSISSDNQLPELKQKSSNSRKKRSTSAOPTVDPDRN 180  
 QY 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDPEKVT 240  
 Db 181 DGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDPEKVT 240  
 QY 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300  
 Db 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300  
 QY 301 SEVHGNAEVSFDDIGGSVSAFSSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360  
 Db 301 SEVHGNAEVSFDDIGGSVSAFSSNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360  
 QY 361 NANIRVYNTGTAPIYVLPITTSVLGKNOFLATIDADENOLSOILAPNNYPPSKNLAPIA 420  
 Db 361 NANIRVYNTGTAPIYVLPITTSVLGKNOFLATIDADENOLSOILAPNNYPPSKNLAPIA 420  
 QY 421 LNAQKFSPTITMNVNQFLEKTKQLRLDQVYGNIAVYVNGRVRVDTGSMSEV 480  
 Db 421 LNAQKFSPTITMNVNQFLEKTKQLRLDQVYGNIAVYVNGRVRVDTGSMSEV 480  
 QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540  
 Db 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540  
 QY 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNKMMILIRDKRPHYDR 600  
 Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNITVLDKIKLNKMMILIRDKRPHYDR 600  
 QY 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660  
 Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYVIEBTEGLKEVINDRY 660  
 QY 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTKNTIINPSENGDTSTNG 720  
 Db 661 DMLNSSLRQDQKTFIDFKKYNDKPLXIYISNPNKVNVAVTKNTIINPSENGDTSTNG 720  
 QY 721 IKKIL---IFSCKG 731  
 Db 721 IKKILKKVVLGKG 734

## RESULT 7

US-08-082-849B-31  
 ; Sequence 31, Application US/08082849B  
 ; Patent No. 5677274  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leppa, Stephen H.  
 ; APPLICANT: Klimpel, Kurt R.  
 ; APPLICANT: Arora, Naveen  
 ; APPLICANT: Singh, Yogendra  
 ; APPLICANT: Nichols, Peter J.  
 ; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
 ; TITLE OF INVENTION: Related Methods  
 ; NUMBER OF SEQUENCES: 35  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, Eighth Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-021-601-12

Query Match 98.1%; Score 3703.5; DB 1; Length 903;  
Best Local Similarity 98.6%; Pred. No. 3.5e-257; Indels 3; Gaps 1;  
Matches 724; Conservative 1; Mismatches 6;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60  
QY 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKSWSTASDPYDFEKT 240  
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKSWSTASDPYDFEKT 240  
QY 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
Db 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKFSPTPTMNYNQFLEBKTQKRLDLDQVYGNATYNFENGVRVDTGNSWSEV 480  
Db 421 LNAQKFSPTPTMNYNQFLEBKTQKRLDLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTKEALKIAGFNPNGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTKEALKIAGFNPNGNL 540  
QY 541 QYQKDIETDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNILIRDKRPHYDR 600  
Db 541 QYQKDIETDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNILIRDKRPHYDR 600  
QY 601 NNIAVGADESVEKAEHREVINSSTEGILLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVEKAEHREVINSSTEGILLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDKGTIFDFKYNQDKLPIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDKGTIFDFKYNQDKLPIYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKIL---IFSKG 731  
Db 721 IKKILKKVLGKG 734

RESULT 5  
US-08-082-849B-12  
Sequence 12, Application US/08082849B  
Patent No. 5677274  
GENERAL INFORMATION:  
APPLICANT: Leppa, Stephen H.  
APPLICANT: Klompel, Kurt R.  
APPLICANT: Aroa, Naveen

APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
TITLE OF INVENTION: Related Methods  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-12

Query Match 98.1%; Score 3703.5; DB 1; Length 903;

Best Local Similarity 98.6%; Pred. No. 3.5e-257; Indels 3; Gaps 1;  
Matches 724; Conservative 1; Mismatches 6;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60  
QY 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDRYTATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKSWSTASDPYDFEKT 240  
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGTLTKYKSPKSWSTASDPYDFEKT 240  
QY 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
Db 241 GRIDKNVSPKARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVKGKQTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKFSPTPTMNYNQFLEBKTQKRLDLDQVYGNATYNFENGVRVDTGNSWSEV 480

QY 661 DMLNISSLRQDKTTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720  
DB 661 DMLNISSLRQDKTTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720  
QY 721 IKKILFSKKGYEIG 735  
DB 721 IKKILFSKKGYEIG 735

RESULT 3  
PCT-US94-01624-4  
; Sequence 4, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW  
; STREET: Steuart Street Tower, 20th Floor, One Market  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: PCT/US94/01624  
; FILING DATE: June 25, 1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-01624-4

Query Match 99.3%; Score 3748; DB 5; Length 735;  
Best Local Similarity 99.5%; Pred. No. 1.7e-260;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMWTSTTGDLSPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMWTSTTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWGFFIKVKSDYETATSDADHVTWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
DB 61 QSAIWGFFIKVKSDYETATSDADHVTWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGTTPVDRDN 180  
DB 121 QRENPTKEGLDFKLYWTDSONKKEVISSDNLQPELKQKSSNSRKRKSTSGTTPVDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVNKRFTFLSPWISNHEKKGTLTKSSPEKWSIASDPSDFEYVT 240  
DB 181 DGIPTDSLEVEGYTVDVNKRFTFLSPWISNHEKKGTLTKSSPEKWSIASDPSDFEYVT 240

QY 241 GRIDKNVSPARHPLVAAYPVIVHVMENIILSKNEQSQNTDSETRTISKNTSRTHT 300  
DB 241 GRIDKNVSPARHPLVAAYPVIVHVMENIILSKNEQSQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLTPTTSLVLGKNOTLATIDADENQLSQIILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYNNVLTPTTSLVLGKNOTLATIDADENQLSQIILAPNNYPSKNLAPIA 420  
QY 421 LNAQKFFSPTITMNTNQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGSKNWEV 480  
DB 421 LNAQKFFSPTITMNTNQFLELEKTKQLRLDQVYGNATYVNFENGVRVDTGSKNWEV 480  
QY 481 LPOIQETTARIIFNGKDLNLVRRRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540  
DB 481 LPOIQETTARIIFNGKDLNLVRRRIAANVPSDPLETTKPDMLKEALKIAGFNEPENGNL 540  
QY 541 QYQKDIITBDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKONILIRDKRPHYDR 600  
DB 541 QYQKDIITBDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNAKONILIRDKRPHYDR 600  
QY 601 NNTAVGADESVMKEAAREVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660  
DB 601 NNTAVGADESVMKEAAREVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDKTTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720  
DB 661 DMLNISSLRQDKTTFIDFKKNDKLPXYISNPYKVNVAVTKENTINPSENGDTSTNG 720  
QY 721 IKKILFSKKGYEIG 735  
DB 721 IKKILFSKKGYEIG 735

RESULT 4  
US-08-021-601-12  
; Sequence 12, Application US/08021601  
; Patent No. 5591631  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Nichols, Peter J.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Needle & Rosenberg, P.C.  
; STREET: 133 Carnegie Way, Suite 400  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/021,601  
; FILING DATE: 19930212  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 1414.057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880



1 EVKQENRLNSESSESSQGLGYFSLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
61 QSAIWSGFTKVKKSDYEYTFATSDADNHVTWVDDQEVINKANSNKIRLEKGRLYQIKIY 120  
61 QSAIWSGFTKVKKSDYEYTFATSDADNHVTWVDDQEVINKANSNKIRLEKGRLYQIKIY 120  
121 QRENTEKGLDFKLYWTDSONKKEVSISSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
121 QRENTEKGLDFKLYWTDSONKKEVSISSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
181 DGIPODSEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240  
181 DGIPODSEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240  
241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNTSRTHT 300  
241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNTSRTHT 300  
301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540  
481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540  
541 QYQKDIETEFDFPDQOSTQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600  
541 QYQKDIETEFDFPDQOSTQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600  
601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

RESULT 2  
US-082-849B-4  
Sequence 4, Application US/08082849B  
Patent No. 5677274

## GENERAL INFORMATION:

APPLICANT: Leppla, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
TITLE OF INVENTION: Related Methods  
NUMBER OF SEQUENCES: 35

## CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-4

Query Match 99.3%; Score 3748; DB 1; Length 735;

Best Local Similarity 99.5%; Pred. No. 1.7e-260;

Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFSLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
QY 61 QSAIWSGFTKVKKSDYEYTFATSDADNHVTWVDDQEVINKANSNKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFTKVKKSDYEYTFATSDADNHVTWVDDQEVINKANSNKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVSISSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
DB 121 QRENTEKGLDFKLYWTDSONKKEVSISSNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
QY 181 DGIPODSEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240  
DB 181 DGIPODSEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNTSRTHT 300  
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRVNTGTAPIYVNLPTTSLVLGKNOTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
DB 421 LNAQKPSSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540  
DB 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540  
QY 541 QYQKDIETEFDFPDQOSTQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600  
DB 541 QYQKDIETEFDFPDQOSTQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRPHYDR 600  
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
DB 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 15.0986 Seconds  
(without alignments)  
2513.152 Million cell updates/sec

Title: US-09-848-909A-11  
Perfect score: 3774  
Sequence: 1 EVKQENLLNSESQGLL.....TSTNGIKLIFSKGVEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 5125971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pcp: \*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pcp: \*  
3: /cgn2\_6/prodata/2/iaa/5A COMB.pcp: \*  
4: /cgn2\_6/prodata/2/iaa/5B COMB.pcp: \*  
5: /cgn2\_6/prodata/2/iaa/5A COMB.pcp: \*  
6: /cgn2\_6/prodata/2/iaa/5B COMB.pcp: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3748	99.3	735	1 US-08-021-601-4	Sequence 4, Appl
2	3748	99.3	735	1 US-08-082-849B-4	Sequence 4, Appl
3	3748	99.3	735	5 PCT-US94-01624-4	Sequence 4, Appl
4	3703.5	98.1	903	1 US-08-021-601-12	Sequence 12, Appl
5	3703.5	98.1	903	1 US-08-082-849B-12	Sequence 12, Appl
6	3703.5	98.1	903	5 PCT-US94-01624-12	Sequence 12, Appl
7	3571	94.6	719	1 US-08-082-849B-31	Sequence 31, Appl
8	3571	94.6	719	5 PCT-US94-01624-31	Sequence 31, Appl
9	1275	33.8	288	4 US-09-273-839A-8	Sequence 8, Appl
10	768	20.3	881	3 US-08-960-780-32	Sequence 32, Appl
11	768	20.3	881	3 US-09-073-898-32	Sequence 32, Appl
12	768	20.3	881	4 US-09-307-106-8	Sequence 8, Appl
13	768	20.3	881	4 US-09-850-351A-32	Sequence 32, Appl
14	766	20.3	884	1 US-08-471-033-5	Sequence 5, Appl
15	766	20.3	884	2 US-08-471-044-5	Sequence 5, Appl
16	766	20.3	884	2 US-08-463-483A-5	Sequence 5, Appl
17	766	20.3	884	2 US-08-471-046A-5	Sequence 5, Appl
18	766	20.3	884	2 US-08-470-566B-5	Sequence 5, Appl
19	766	20.3	884	2 US-08-469-334-5	Sequence 5, Appl
20	766	20.3	884	3 US-09-300-529-5	Sequence 5, Appl
21	766	20.3	1346	1 US-08-471-033-23	Sequence 23, Appl
22	766	20.3	1346	2 US-08-471-044-23	Sequence 23, Appl
23	766	20.3	1346	2 US-08-463-483A-23	Sequence 23, Appl
24	766	20.3	1346	2 US-08-471-046A-23	Sequence 23, Appl
25	766	20.3	1346	2 US-08-470-566B-23	Sequence 23, Appl
26	766	20.3	1346	2 US-08-469-334-23	Sequence 23, Appl
27	766	20.3	1346	3 US-09-300-529-23	Sequence 23, Appl

28	765	20.3	852	1 US-08-471-033-36	Sequence 36, Appl
29	765	20.3	852	2 US-08-471-044-36	Sequence 36, Appl
30	765	20.3	852	2 US-08-463-483A-36	Sequence 36, Appl
31	765	20.3	852	2 US-08-471-046A-36	Sequence 36, Appl
32	765	20.3	852	2 US-08-470-566B-36	Sequence 36, Appl
33	765	20.3	852	3 US-09-300-529-36	Sequence 36, Appl
34	765	20.3	852	3 US-08-469-334-36	Sequence 36, Appl
35	765	20.3	1338	1 US-08-471-033-50	Sequence 50, Appl
36	765	20.3	1338	2 US-08-471-044-50	Sequence 50, Appl
37	765	20.3	1338	2 US-08-463-483A-50	Sequence 50, Appl
38	765	20.3	1338	2 US-08-471-046A-50	Sequence 50, Appl
39	765	20.3	1338	2 US-08-470-566B-50	Sequence 50, Appl
40	765	20.3	1338	3 US-08-469-334-50	Sequence 50, Appl
41	765	20.3	1338	3 US-09-300-529-50	Sequence 50, Appl
42	755.5	20.0	784	3 US-09-371-913A-7	Sequence 7, Appl
43	755.5	20.0	784	4 US-09-967-803-7	Sequence 7, Appl
44	746.5	19.8	860	4 US-09-307-106-48	Sequence 48, Appl
45	736.5	19.5	834	1 US-08-471-033-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1  
US-08-021-601-4  
; Sequence 4, Application US/08021601  
; Patent No. 5591631  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Nichols, Peter J.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Needle & Rosenberg, P.C.  
; STREET: 133 Carnegie Way, Suite 400  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/021,601  
; FILING DATE: 19930212  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/POCKET NUMBER: 1414.057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-9770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-021-601-4

Query Match 99.3%; Score 3748; DB 1; Length 735;

Best Local Similarity 99.5%; Pred. No. 1.7e-260;

Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENLLNSESQGLLGYFSDNLFQAPMVVTSSTTCDDLSIPSELENISENYQF 60

DB 690 DMLNSSLRODQKTFIDFKYNDKLPYISNPYKVVAVTKENTIIINPSENGDTSTNG 749

QY 721 IKKILIFSKKGYEIG 735

DB 750 IKKILIFSKKGYEIG 764

RESULT 15

AAB47306

ID AAB47306 standard; protein; 764 AA.

AC AAB47306;

DT 29-AUG-2001 (first entry)

DE Wild type B. anthracis protective antigen.

XX Lethal factor; LF; immunogen; LF4; protective antigen; PA; DNA vaccine;

XX humoral; cell-mediated; immune memory response.

XX Bacillus anthracis.

XX Key Location/Qualifiers

FT Peptide 1..29

FT /label= signal peptide

FT /note= "Not given in the specification"

FT Protein 30..764

FT /label= PA

FT Peptide 204..764

FT /label= pCPA

XX WO2000145639-A2.

PN 28-JUN-2001.

XX 21-DEC-2000; 2000WO-US034912.

PF 22-DEC-1999; 99US-0171459P.

PR (OHIS ) UNIV OHIO STATE RES FOUND.

XX (GALL/) GALLOWAY D R.

PA (MATE/) MATECZUN A J.

XX Galloway DR, Mateczun AJ;

XX WPI; 2001-408540/43.

DR N-PSDB; AAC86016.

XX Protecting animal against lethal infection with Bacillus anthracis, by

PT administering wildtype or mutated form of Bacillus anthracis lethal

PT factor protein or its fragment or a nucleic acid encoding the mutated

PT protein.

XX Claim 5; Fig 2; 33pp; English.

XX This sequence shows the B. anthracis protective antigen (PA). An

CC immunogenic fragment of PA, pCPA, can be used to produce an immune

CC response which protects an animal against lethal infection with Bacillus

CC anthracis. DNA encoding the B. anthracis PA can be used in conjunction

CC with DNA encoding the lethal factor (LF) in a DNA vaccine. Using a DNA

CC vaccine which encodes the mutated LF protein or fragment alone or in

CC combination with a DNA encoding the PA protein or its fragment, both

CC components (humoral and cell-mediated) of the immune system are

CC stimulated, which results in longer term immune memory response. The

CC combined use of a mutated LF and PA gene or their fragments results in a

CC higher level of immune response as judged by overall serum antibody in

CC titers for LF and PA antigens, than the use of either LF or PA genes in

CC separate immunizations

XX Sequence 764 AA;

SQ

Query Match 99.3%; Score 3748; DB 4; Length 764;

Best Local Similarity 99.5%; Pred. No. 1.3e-239;

Matches	731;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
QY	1	EVKQENRLNSESSESSQGLGYPFDLNFQAPMVVTSTTTGDLSPSSSELENIPSENQYF	60						
DB	30	EVKQENRLNSESSESSQGLGYPFDLNFQAPMVVTSTTTGDLSPSSSELENIPSENQYF	89						
QY	61	QSAIWSGFYKVKSDXENTFATSADNHVTMVDDQEVINKASNNKIRLEKGRLYQIKIY	120						
DB	90	QSAIWSGFYKVKSDXENTFATSADNHVTMVDDQEVINKASNNKIRLEKGRLYQIKIY	149						
QY	121	QRENPTKGLDFKLYWTDSONKKEVVISDNLQLPKQKSSNSRKKKSTSGAGTVPDRDN	180						
DB	150	QRENPTKGLDFKLYWTDSONKKEVVISDNLQLPKQKSSNSRKKKSTSGAGTVPDRDN	209						
QY	181	DGIPDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWTASDPYSDPEKVT	240						
DB	210	DGIPDSLEVEGYVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWTASDPYSDPEKVT	269						
QY	241	GRIDKNVSPPEARHPLVAAAYPIVHVDMENIILSKNEDQSTQNTDSETTISKTSTSTHT	300						
DB	270	GRIDKNVSPPEARHPLVAAAYPIVHVDMENIILSKNEDQSTQNTDSETTISKTSTSTHT	329						
QY	301	SEVHGNAEVAHAFDPDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL	360						
DB	330	SEVHGNAEVAHAFDPDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEWTGLNTADTARL	389						
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSQLAPNNYYPKKNLAPIA	420						
DB	390	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIDADENOLSQLAPNNYYPKKNLAPIA	449						
QY	421	LNAOKKFSSTPITMYNQFLEKTKQLRLDQVYGNATYFENFENGRVVDVTCNNWSEV	480						
DB	450	LNAOKKFSSTPITMYNQFLEKTKQLRLDQVYGNATYFENFENGRVVDVTCNNWSEV	509						
QY	481	LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAPGNEPNGL	540						
DB	510	LPQIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAPGNEPNGL	569						
QY	541	OYQKGDITEFDNFDOQTSONIKQLAELNATNTYVLDKIKNKXNLIIRDKRFHYDR	600						
DB	570	OYQKGDITEFDNFDOQTSONIKQLAELNATNTYVLDKIKNKXNLIIRDKRFHYDR	629						
QY	601	NNIAYGADESUVKEAHREVINSSTEGLLNLDKIRKILSGYIIVEIDTGLKEVINDRY	660						
DB	630	NNIAYGADESUVKEAHREVINSSTEGLLNLDKIRKILSGYIIVEIDTGLKEVINDRY	689						
QY	661	DMLNSSLRQDGKTFIDFKYNDKLPYISNPYKVVAVTKENTIIINPSENGDTSTNG	720						
DB	690	DMLNSSLRQDGKTFIDFKYNDKLPYISNPYKVVAVTKENTIIINPSENGDTSTNG	749						
QY	721	IKKILIFSKKGYEIG 735							
DB	750	IKKILIFSKKGYEIG 764							

Search completed: May 3, 2004, 19:36:00

Job time : 51.1208 secs

```
QY 121 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPKQKSSNRKKESTAGTVPDRDN 180
DB 149 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPKQKSSNRKKESTAGTVPDRDN 208
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSASTDPYSDFEKT 240
DB 209 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSASTDPYSDFEKT 268
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
DB 269 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 328
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 329 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 388
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNTQTLATIDADENQLSLAPNYYPSKNLAPIA 420
DB 389 NANIRYVNTGTAPIYVNLPTTSLVLGKNTQTLATIDADENQLSLAPNYYPSKNLAPIA 448
QY 421 LNAOKKRSSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480
DB 449 LNAQDDFSSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 508
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 540
DB 509 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 568
QY 541 QYQKDIITEFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRPFYDR 600
DB 569 QYQKDIITEFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRPFYDR 628
QY 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
DB 629 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 688
QY 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVAVTKNTIINPSENGDSTNG 720
DB 689 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVAVTKNTIINPSENGDSTNG 748
QY 721 IKKILIFSKKGYEIG 735
DB 749 IKKILIFSKKGYEIG 763

RESULT 14
AA56958
ID AA56958 standard; protein; 764 AA.
XX
AC AA56958;
XX
DT 25-APR-2000 (first entry)
XX
DE B. anthracis protective antigen (PA) protein.
XX
KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;
KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.
XX
OS Bacillus anthracis.
XX
PN WO200002522-A2.
XX
PD 20-JAN-2000.
XX
PF 09-JUL-1999; 99WO-US015568.
XX
PR 10-JUL-1998; 98US-0092416P.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;
XX
```

```
DR WPI; 2000-182165/16.
DR N-PSDB; AA256874.
XX
PT Recombinant DNA construct useful as vaccines for anthrax, in producing
PT host cells for analyzing the drugs and agents inhibiting anthrax.
XX
PS Disclosure; Page 33; 35pp; English.
XX
CC The invention provides a recombinant DNA construct that comprises a
CC vector and at least one nucleic acid (or its fragment) encoding a
CC combination of Bacillus anthracis proteins, selected from protective
CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA
CC with its secretory signals replaced with those of tissue plasminogen
CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine
CC for anthrax and in producing infectious alpha virus particles. These
CC particles, expressing the B. anthracis proteins are useful also as
CC vaccines for anthrax. Host cells transformed with the construct are
CC useful for analyzing the effectiveness of drugs and agents that inhibit
CC anthrax or B. anthracis proteins. The present sequence represents a B.
CC anthracis PA protein
XX
SQ Sequence 764 AA;
Query Match 99.3%; Score 3748; DB 3; Length 764;
Best Local Similarity 99.5%; Pred. No. 1.3e-239;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 EVKQENRLNSESSESSQGLLYFSDNLFQAPMVVTSSTGDLSPSSSELENIQSENQYF 60
DB 30 EVKQENRLNSESSESSQGLLYFSDNLFQAPMVVTSSTGDLSPSSSELENIQSENQYF 89
QY 61 QSAIWSGFTKVKKSDBYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 90 QSAIWSGFTKVKKSDBYTFATGADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 149
QY 121 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPKQKSSNRKKESTAGTVPDRDN 180
DB 150 QRENPTKGLDFKLYWTDSONKKEVISNDNLQLPKQKSSNRKKESTAGTVPDRDN 209
QY 181 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSASTDPYSDFEKT 240
DB 210 DGIPLDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSASTDPYSDFEKT 269
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 300
DB 270 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDOSTQNTDSETRTISKNTSTSRHT 329
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 330 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 389
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNTQTLATIDADENQLSLAPNYYPSKNLAPIA 420
DB 390 NANIRYVNTGTAPIYVNLPTTSLVLGKNTQTLATIDADENQLSLAPNYYPSKNLAPIA 449
QY 421 LNAOKKRSSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 480
DB 450 LNAQDDFSSTPTIMNYNOFLEKTKQLRLDQVYGNIAFYNGRVRVDTGNSWSEV 509
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 540
DB 510 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPGFNEPNGNL 569
QY 541 QYQKDIITEFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRPFYDR 600
DB 570 QYQKDIITEFDNFQOQTSQNIKNQLAELNATNIYTVLDKIKLAKNNILIRDKRPFYDR 629
QY 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 660
DB 630 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIYEIEDTEGLKEVINDRY 689
QY 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNYKVNVAVTKNTIINPSENGDSTNG 720
```



Db 541 QYQKIDTEFDNFQDQTSQNKQLAELNATNYTVLDKIKLAKNNILIRDRFHYDR 600  
QY 601 NNIAVGADSVVKEAAREVINSSTGLINIDKDKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADSVVKEAAREVINSSTGLINIDKDKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDQKTFIDFKYNDKULPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRQDQKTFIDFKYNDKULPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

RESULT 11  
AAM51488  
ID AAM51488 standard; protein; 735 AA.  
AC AAM51488;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Anthrax PA mutant D425A.  
XX  
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
OS Bacillus anthracis.  
XX Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 425 /note= "Wild-type Asp substituted by Ala"  
XX  
PN WO200182788-A2.  
XX  
PD 08-NOV-2001.  
XX  
PF 04-MAY-2001; 2001WO-US014372.  
XX  
PR 04-MAY-2000; 2000US-0201800P.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX  
PT Protecting humans against anthrax using mutant B groups (anthrax  
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
PT anthracis.  
XX  
PS Claim 4; Page; 77pp; English.  
XX  
CC The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;

Query Match 99.3%; Score 3748; DB 5; Length 735;  
Best Local Similarity 99.5%; Pred. No. 1.2e-239;  
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLLYFSDLNFAQPMVVTSSTTGDLSPSSSELENIPSENQYF 60  
61 QSAIWGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120  
61 QSAIWGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIQY 120  
121 QRENPTKGLDFKLYWTDSONKEVSSDNQLPELKQKSSNSRKKSTAGPTVPDRN 180  
121 QRENPTKGLDFKLYWTDSONKEVSSDNQLPELKQKSSNSRKKSTAGPTVPDRN 180  
181 DGIPIPSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKASTASDPYDSEKVT 240  
181 DGIPIPSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKASTASDPYDSEKVT 240  
241 GRIDKNVSPARPLVAAPYIVHVDMENTILSKVEDQSTQNTDSETRTTSKNTSTRTHT 300  
241 GRIDKNVSPARPLVAAPYIVHVDMENTILSKVEDQSTQNTDSETRTTSKNTSTRTHT 300  
301 SEVHGNAEVSHPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITMGLTADTARL 360  
301 SEVHGNAEVSHPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITMGLTADTARL 360  
361 NANIRYVNTGTAPIYVLPPTSLVGLKQNOTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420  
361 NANIRYVNTGTAPIYVLPPTSLVGLKQNOTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420  
421 LNAQKFFSSPTITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNNWSEV 480  
421 LNAQKFFSSPTITMNYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGNNWSEV 480  
481 LPOIQTETARIIFNGKDLNIVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENG 540  
481 LPOIQTETARIIFNGKDLNIVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENG 540  
541 QYQKIDTEFDNFQDQTSQNKQLAELNATNYTVLDKIKLAKNNILIRDRFHYDR 600  
541 QYQKIDTEFDNFQDQTSQNKQLAELNATNYTVLDKIKLAKNNILIRDRFHYDR 600  
601 NNIAVGADSVVKEAAREVINSSTGLINIDKDKILSGYIVIEIDTEGLKEVINDRY 660  
601 NNIAVGADSVVKEAAREVINSSTGLINIDKDKILSGYIVIEIDTEGLKEVINDRY 660  
661 DMLNSSLRQDQKTFIDFKYNDKULPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
661 DMLNSSLRQDQKTFIDFKYNDKULPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
721 IKKILIFSKKGYEIG 735  
721 IKKILIFSKKGYEIG 735

RESULT 12  
AAY56959  
ID AAY56959 standard; protein; 736 AA.  
XX  
AC AAY56959;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE B. anthracis MAT-PA protein.  
XX  
KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;  
XX tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.  
XX  
OS Bacillus anthracis.  
XX  
PN WO200002522-A2.  
XX  
PD 20-JAN-2000.  
XX  
PF 09-JUL-1999; 99WO-US015568.  
XX  
PR 10-JUL-1998; 98US-0092416P.

```
SQ Sequence 735 AA;
Query Match 99.3%; Score 3748; DB 2; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.2e-239;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKSTAGPVPDRN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKSTAGPVPDRN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240
QY 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVGLKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVYNTGTAPIYVNLPTTSLVGLKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480
DB 421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 601 NNIAVGADSVVKEAHRVINSSTGLMINTDKDIRKILSGYVIEIDTEGLKEVINDRY 660
DB 601 NNIAVGADSVVKEAHRVINSSTGLMINTDKDIRKILSGYVIEIDTEGLKEVINDRY 660
QY 661 DMLNLSLQDQKTFIDPKKYNKLPXYISNPNYKVNYYAVTKENTIINPSENGDTSNG 720
DB 661 DMLNLSLQDQKTFIDPKKYNKLPXYISNPNYKVNYYAVTKENTIINPSENGDTSNG 720
QY 721 IKKILIFSCKGYEIG 735
DB 721 IKKILIFSCKGYEIG 735

RESULT 10
AAMS1483
ID AAMS1483 standard; protein; 735 AA.
XX
AC AAMS1483;
XX
XX 01-FEB-2002 (first entry)
XX
XX Anthrax PA protein.
XX
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.
XX
XX Bacillus anthracis.
```

```
XX WO200182788-A2.
XX
XX 08-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014372.
XX
XX 04-MAY-2000; 2000US-0201800P.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Collier RJ, Sellman BR;
XX
XX MPI; 2002-017725/02.
XX
XX N-PSDB; AA:99904.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX
XX Disclosure; Fig 13; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. The present sequence is that of the
XX anthrax PA protein
XX
XX Sequence 735 AA;
Query Match 99.3%; Score 3748; DB 5; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.2e-239;
Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENYF 60
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKSDSEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKSTAGPVPDRN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPDLKQKSSNSRKKSTAGPVPDRN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEYV 240
QY 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
DB 241 GRIDKNVSPERHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVGLKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVYNTGTAPIYVNLPTTSLVGLKNQTLATIDADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480
DB 421 LNAQKFFSTPTITMNYNQFLEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIITEFDNFDOQTSONIKNLAELNATNIYTVLDDKILNAKNNILIRDKRPHYDR 600
```

KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
 OS Bacillus anthracis.  
 XX Synthetic.  
 FH Key Location/Qualifiers  
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Asn"  
 XX WO2000182788-A2.  
 XX PD 08-NOV-2001.  
 XX PF 04-MAY-2001; 2001WO-US014372.  
 XX PR 04-MAY-2000; 2000US-0201800P.  
 XX PA (HARD ) HARVARD COLLEGE.  
 XX PI Collier RJ, Sellman BR;  
 XX WPI; 2002-017725/02.  
 XX PT Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.  
 XX Claim 4; Page; 77pp; English.  
 XX The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
 XX Sequence 735 AA;  
 SQ

Query Match 99.3%; Score 3749; DB 5; Length 735;  
 Best Local Similarity 99.5%; Pred. No. 1.1e-239;  
 Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EVKQENRLINSESSQGLGYRFDLNFQAPMVVTSSTTGLSIPSELENIPSENOYF 60  
 DB 1 EVKQENRLINSESSQGLGYRFDLNFQAPMVVTSSTTGLSIPSELENIPSENOYF 60  
 QY 61 QSAIWSGPIKVKSDYTFATSAADNHVTMWYDDQEVINKASNSKIRLEKGRLYQIKIY 120  
 DB 61 QSAIWSGPIKVKSDYTFATSAADNHVTMWYDDQEVINKASNSKIRLEKGRLYQIKIY 120  
 QY 121 QRENTEKGLDPLKLYWTDSONKKEVISSDNQLQLPELKQSSNRKKSSTSGAPVPRDN 180  
 DB 121 QRENTEKGLDPLKLYWTDSONKKEVISSDNQLQLPELKQSSNRKKSSTSGAPVPRDN 180  
 QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIEHKGLTKYKSSPEKWSSTASDPYDFEKVT 240  
 DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIEHKGLTKYKSSPEKWSSTASDPYDFEKVT 240  
 QY 241 GRIDKNVSPERHPVLAAPVPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSTHT 300  
 DB 241 GRIDKNVSPERHPVLAAPVPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSTHT 300  
 QY 301 SEVHGNAEVHASFFDIGSVSAGFSNSSTVAIDHLSLAGERTWAEITMGLNTADTARL 360  
 DB 301 SEVHGNAEVHASFFDIGSVSAGFSNSSTVAIDHLSLAGERTWAEITMGLNTADTARL 360  
 QY 361 NANIRYVNTGAPINNVPTTSLVLGKQKQTLATDADENQISQILAPNNYPSKXNLAPIA 420  
 DB 361 NANIRYVNTGAPINNVPTTSLVLGKQKQTLATIKAKENQISQILAPNNYPSKXNLAPIA 420

QY 421 LNAQKFSSTPTIMYNOFLELEKTKQLRLDQVYGNIAATYNFENGVRVVDTSNWSSEV 480  
 DB 421 LNAQKFSSTPTIMYNOFLELEKTKQLRLDQVYGNIAATYNFENGVRVVDTSNWSSEV 480  
 QY 481 LPQIQETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
 DB 481 LPQIQETTARIIFNGKDLNVERRIAANPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
 QY 541 QYQKDKITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKEFHVDYR 600  
 DB 541 QYQKDKITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKEFHVDYR 600  
 QY 601 NNIAGADESVVKEAHREVINSSTGLLINDKIDRKILSGYIVIEDETEGLKEVINDRY 660  
 DB 601 NNIAGADESVVKEAHREVINSSTGLLINDKIDRKILSGYIVIEDETEGLKEVINDRY 660  
 QY 661 DMNLSLRODQKTFIDFKKYNKPLPYISNPNYKVVYAVTKENTINPSENGDTSTNG 720  
 DB 661 DMNLSLRODQKTFIDFKKYNKPLPYISNPNYKVVYAVTKENTINPSENGDTSTNG 720  
 QY 721 IKKILIFSKKGYEIG 735  
 DB 721 IKKILIFSKKGYEIG 735

RESULT 9  
 AAR60179  
 ID AAR60179 standard; protein; 735 AA.  
 XX  
 AC AAR60179;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 03-APR-1995 (first entry)  
 DE Protective antigen of Bacillus anthracis.  
 XX Anthrax; Bacillus anthracis; fusion protein; protective antigen;  
 KW protective antigen; cell killing; targeting; targeting; pathogen;  
 KW intracellular; HIV; human immunodeficiency virus; toxin.  
 XX Bacillus anthracis.  
 OS  
 XX WO9418332-A2.  
 PN  
 PD 18-AUG-1994.  
 XX 14-FEB-1994; 94WO-US001624.  
 XX 12-FEB-1993; 93US-00021601.  
 PR 25-JUN-1993; 93US-00082849.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Leppla SH, Klimpel K, Arora N, Singh Y, Nichols PJ;  
 XX WPI; 1994-279753/34.  
 DR N-PSDB; AAQ70180.  
 XX  
 PT Nucleic acid encoding anthrax toxin fusion protein - useful for  
 PT targeting toxin to specific cells, eg for killing tumour cells or HIV-  
 PT infected cells.  
 XX  
 PS Disclosure; Page 81-83; 124pp; English.  
 CC  
 CC The sequence encoding the protective antigen of Bacillus anthracis may be  
 CC used in the construction of a nucleic acid which encodes a fusion protein  
 CC comprising the anthrax protective antigen binding domain of the native  
 CC anthrax lethal factor and a sequence encoding an activity inducing domain  
 CC of a second protein. The fusion proteins are useful for the specific  
 CC killing of tumour cells or the killing of cells infected with  
 CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to  
 CC correct PN field.)  
 XX



Db 361 NANIRVNTGTAPIYVNLPTTSLVGLKQTLATIKAKENQLSQLAPNYYPSKXNLAPIA 420  
 Qy 421 LNAOKFSSPTPTMKNQFLELEKTKQLRLDQVYGNATYNGRVRVDTGSNWEV 480  
 Db 421 LNAQDFSSPTPTMKNQFLELEKTKQLRLDQVYGNATYNGRVRVDTGSNWEV 480  
 Qy 481 LPQIQETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIARFNEPNGNL 540  
 Db 481 LPQIQETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIARFNEPNGNL 540  
 Qy 541 QYQKDIETDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600  
 Db 541 QYQKDIETDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600  
 Qy 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
 Db 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
 Qy 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
 Db 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
 Qy 721 IKKILIFSCKGYEIG 735  
 Db 721 IKKILIFSCKGYEIG 735

RESULT 7  
 AAMS1487  
 ID AAMS1487 standard; protein; 735 AA.  
 AC AAMS1487;  
 XX

DT 01-FEB-2002 (first entry)  
 DE Anthrax PA mutant K397Q.  
 XX

KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 KM B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
 XX  
 OS Bacillus anthracis.  
 XX  
 XX Synthetic.  
 XX

FH Key Location/Qualifiers  
 FT Misc-difference 397  
 XX

FT /note= "Wild-type Lys substituted by Gln"  
 XX

PN WC200182788-A2.  
 XX

XX 08-NOV-2001.  
 PD

XX 04-MAY-2001; 2001WO-US014372.  
 XX

XX 04-MAY-2000; 2000US-0201800P.  
 PR

XX (HARD ) HARVARD COLLEGE.  
 PA

XX Collier RJ, Sellman BR;  
 PI

XX WPI; 2002-017725/02.  
 DR

XX Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.  
 XX

PS Claim 4; Page; 77pp; English.  
 XX

CC The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC

CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAMS1483)  
 XX  
 SQ Sequence 735 AA;

Query Match 99.3%; Score 3749; DB 5; Length 735;

Best Local Similarity 99.5%; Pred. No. 1.1e-239;

Matches 731; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESESSQGLLYFFDLNFOAPMWVTSSTGDLSPSSSELENIPSENOYF 60

Db 1 EVKQENRLNESESSQGLLYFFDLNFOAPMWVTSSTGDLSPSSSELENIPSENOYF 60

Qy 61 QSAIWGFIKVKKSDBYTATADNHVTWVDDQEVINKASNSKIRLSEKGLYQIKIY 120

Db 61 QSAIWGFIKVKKSDBYTATADNHVTWVDDQEVINKASNSKIRLSEKGLYQIKIY 120

Qy 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQSSNSRKRSTASGTVDPDRN 180

Db 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQSSNSRKRSTASGTVDPDRN 180

Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDPEKVT 240

Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDPEKVT 240

Qy 241 GRIDKNVSPARHPIVAAPIVHVDMENILSKNEDOSTCNTDSETRTISKNTSRTHT 300

Db 241 GRIDKNVSPARHPIVAAPIVHVDMENILSKNEDOSTCNTDSETRTISKNTSRTHT 300

Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERWAEWTGMLNTADTASL 360

Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERWAEWTGMLNTADTASL 360

Qy 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQTLATIKAKENQLSQLAPNYYPSKXNLAPIA 420

Db 361 NAMIRYVNTGTAPIYVNLPTTSLVGLKQTLATIKAKENQLSQLAPNYYPSKXNLAPIA 420

Qy 421 LNAQKFPSSPTPTMKNQFLELEKTKQLRLDQVYGNATYNGRVRVDTGSNWEV 480

Db 421 LNAQKFPSSPTPTMKNQFLELEKTKQLRLDQVYGNATYNGRVRVDTGSNWEV 480

Qy 481 LPQIQETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIARFNEPNGNL 540

Db 481 LPQIQETARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKEALKIARFNEPNGNL 540

Qy 541 QYQKDIETDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600

Db 541 QYQKDIETDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKXNMLIRDKRFHYDR 600

Qy 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660

Db 601 NNIAVGADES SVKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660

Qy 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

Db 661 DMLNSSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

Qy 721 IKKILIFSCKGYEIG 735

Db 721 IKKILIFSCKGYEIG 735

RESULT 8  
 AAMS1489  
 ID AAMS1489 standard; protein; 735 AA.  
 XX  
 XX AAMS1489;  
 AC

XX 01-FEB-2002 (first entry)  
 DT

XX Anthrax PA mutant D425N.  
 XX

CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;

Query Match 99.4%; Score 3753; DB 5; Length 735;  
Best Local Similarity 99.6%; Pred. No. 5.8e-240;  
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSESSQGLLYGFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSESSQGLLYGFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKQY 120  
DB 61 QSAIWSGFIKVKSDDEYTFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRKSTSGPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPYDFEKT 240  
DB 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSOILLAPNNYPSKNLAPIA 420  
DB 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSOILLAPNNYPSKNLAPIA 420  
QY 421 LNAQKFSSTPTIMNYNQFLEKTKQLRLDTPDQVYGNATYFNFGVRVVDTSNWSSEV 480  
DB 421 LNAQKFSSTPTIMNYNQFLEKTKQLRLDTPDQVYGNATYFNFGVRVVDTSNWSSEV 480  
QY 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTPDMTLKALKIAGFNEPGNL 540  
DB 481 LPQIQETARIIFNGKDLNVERRIAANVPSDPLETTPDMTLKALKIAGFNEPGNL 540  
QY 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 600  
DB 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNAMNILLIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLINIDKDIRKILSGYVIEDETEGLKEVINDRY 660  
DB 601 NNIAVGADESUVKEAHEVINSSTEGLLINIDKDIRKILSGYVIEDETEGLKEVINDRY 660  
QY 661 DMLNISLQDQKGTFFDFKKYNDKLPYISNPNYKNVYAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNISLQDQKGTFFDFKKYNDKLPYISNPNYKNVYAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 6  
AAM51490  
ID AAM51490 standard; protein; 735 AA.  
XX  
AC AAM51490;  
XX  
DT 01-FEB-2002 (first entry)

XX Anthrax PA mutant D425E.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.  
XX Bacillus anthracis.  
OS Synthetic.  
XX Key Location/Qualifiers  
FT Misc-difference 425 /note= "Wild-type Asp substituted by Glu"  
XX WO200182788-A2.  
XX 08-NOV-2001.  
XX 04-MAY-2001; 2001WO-US014372.  
XX 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX Claim 4; Page; 77pp; English.  
XX The invention relates to antibacterial agents comprising mutant forms of  
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX moiety is anthrax protective antigen (PA) and using these mutants or  
XX compositions of them for protecting against Bacillus anthracis infections  
XX in humans, especially as vaccines. Note: The present sequence is not  
XX given in the specification but is derived from the Bacillus anthracis  
XX wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;

Query Match 99.4%; Score 3750; DB 5; Length 735;  
Best Local Similarity 99.5%; Pred. No. 9.1e-240;  
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSESSQGLLYGFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSESSQGLLYGFFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKQY 120  
DB 61 QSAIWSGFIKVKSDDEYTFATADNHVTWVDDQVINKASNSKIRLEKGRLYQIKQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRKSTSGPTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVTSNDNLQLPKQKSSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPYDFEKT 240  
DB 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENTILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIDADENQLSOILLAPNNYPSKNLAPIA 420



QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQSLQILAPNNYPSKNLAPIA 420  
 |||||  
 DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQSLQILAPNNYPSKNLAPIA 420  
 |||||  
 QY 421 LNAQKDFSTPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNFENGVRVDTGNNWSEV 480  
 |||||  
 DB 421 LNAQKDFSTPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNFENGVRVDTGNNWSEV 480  
 |||||  
 QY 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPKNL 540  
 |||||  
 DB 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPKNL 540  
 |||||  
 QY 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKNNILIRDRKRFHYDR 600  
 |||||  
 DB 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKNNILIRDRKRFHYDR 600  
 |||||  
 QY 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
 |||||  
 DB 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
 |||||  
 QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLXISPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
 |||||  
 DB 661 DMLNISSLRQDGKTFIDFKKYNKDLPLXISPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
 |||||  
 QY 721 IKKILIFSCKGYEIG 735  
 |||||  
 DB 721 IKKILIFSCKGYEIG 735  
 |||||

## RESULT 3

AAMS1485  
 ID AAMS1485 standard; protein; 735 AA.

XX AC AAMS1485;  
 XX

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant K397D.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.

XX Bacillus anthracis.

OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"

XX WC200182788-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD ) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

XX Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B

CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAMS1483)

XX SQ Sequence 735 AA;

Query Match 99.5%; Score 3755; DB 5; Length 735;

Best Local Similarity 99.6%; Pred. No. 4.2e-240;

Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLGYYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSNQYF 60

DB 1 EVKQENRLNSESSESSQGLLGYYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSNQYF 60

QY 61 QSAIWGFIKVKSDSYTFATSNADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120

DB 61 QSAIWGFIKVKSDSYTFATSNADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRKSTAGTVPDRDN 180

DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQKSSNSRKRKSTAGTVPDRDN 180

QY 181 DGIPDSLEVEGYVDVKNKRTFLSPMISNHEKGLTKVKSSEKASTADPYSDPEKVT 240

DB 181 DGIPDSLEVEGYVDVKNKRTFLSPMISNHEKGLTKVKSSEKASTADPYSDPEKVT 240

QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300

DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300

QY 301 SEVHGNAEVHASFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

DB 301 SEVHGNAEVHASFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQSLQILAPNNYPSKNLAPIA 420

DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIDADENQSLQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSTPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNFENGVRVDTGNNWSEV 480

DB 421 LNAQKDFSTPTITMNYNOFLELEKTKQLRLDQVYGNIAIYNFENGVRVDTGNNWSEV 480

QY 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPKNL 540

DB 481 LPQIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGFNEPKNL 540

QY 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKNNILIRDRKRFHYDR 600

DB 541 QYQKDIETFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLAKNNILIRDRKRFHYDR 600

QY 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660

DB 601 NNIAVGADESUVKEAHEVINSSTGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660

QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLXISPNYKVNVAVTKENTIIINPSENGDTSTNG 720

DB 661 DMLNISSLRQDGKTFIDFKKYNKDLPLXISPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILIFSCKGYEIG 735

DB 721 IKKILIFSCKGYEIG 735

## RESULT 4

AAMS1491

ID AAMS1491 standard; protein; 735 AA.

XX AC AAMS1491;

XX 01-FEB-2002 (first entry)

XX

CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. Note: The present sequence is not  
CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;  
  
Query Match 100.0%; Score 3774; DB 5; Length 735;  
Best Local Similarity 100.0%; Pred. No. 2.3e-241; Indels 0; Gaps 0;  
Matches 735; Conservative 0; Mismatches 0;  
  
QY 1 EVKQENRLNSESSESSQGLLGYFFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLGYFFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60  
  
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
  
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180  
  
QY 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKKGITKYSPEKSTASDPYDFEKT 240  
Db 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKKGITKYSPEKSTASDPYDFEKT 240  
  
QY 241 GRIDKNVSPARHPLVAAPYVHVDMENILSKNEQSTQNTDSETRTSKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAPYVHVDMENILSKNEQSTQNTDSETRTSKNTSTSRHT 300  
  
QY 301 SEVHGNAEVHAFDIOGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAFDIOGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
  
QY 361 NANIRVNTGTPIYVNLSTSLVGNKOTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTPIYVNLSTSLVGNKOTLATIDADENQLSQILAPNNYPSKNLAPIA 420  
  
QY 421 LNAQKFSSTPIYVNLSTSLVGNKOTLATIDADENQLSQILAPNNYPSKNLAPIA 480  
Db 421 LNAQKFSSTPIYVNLSTSLVGNKOTLATIDADENQLSQILAPNNYPSKNLAPIA 480  
  
QY 481 LPOIQETARIIFNGKDLNVERRIAANVPSDPLETTKEDMTLKEALKIAGFENPGNL 540  
Db 481 LPOIQETARIIFNGKDLNVERRIAANVPSDPLETTKEDMTLKEALKIAGFENPGNL 540  
  
QY 541 QYQKIDITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKMILLIRDKRFHYDR 600  
Db 541 QYQKIDITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNAKMILLIRDKRFHYDR 600  
  
QY 601 NNTAVGADESVMKAEHREVINSSTEGLLNIDKDIRKILSGYVIEDTEGLKEVINDRY 660  
Db 601 NNTAVGADESVMKAEHREVINSSTEGLLNIDKDIRKILSGYVIEDTEGLKEVINDRY 660  
  
QY 661 DMLNISLRDQGTFFIDFKYNDKLPYISNPNYKVVYAVTKENTIIINPSENGTSTING 720  
Db 661 DMLNISLRDQGTFFIDFKYNDKLPYISNPNYKVVYAVTKENTIIINPSENGTSTING 720  
  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735  
  
RESULT 2  
AAM51493  
ID AAM51493 standard; protein; 735 AA.  
XX  
AC AAM51493;  
XX  
DT 01-FEB-2002 (first entry)

XX Anthrax PA mutant K397D/D425K.  
DE  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX  
OS Bacillus anthracis.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"  
FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"  
XX  
XX WO200182788-A2.  
XX  
XX 08-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US014372.  
XX  
XX 04-MAY-2000; 2000US-0201800P.  
XX  
XX (HARD ) HARVARD COLLEGE.  
XX  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX  
XX Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX  
XX Claim 4; Page; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of  
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX moiety is anthrax protective antigen (PA) and using these mutants or  
XX compositions of them for protecting against Bacillus anthracis infections  
XX in humans, especially as vaccines. Note: The present sequence is not  
XX given in the specification but is derived from the Bacillus anthracis  
XX wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;  
  
Query Match 99.7%; Score 3761; DB 5; Length 735;  
Best Local Similarity 99.7%; Pred. No. 1.7e-240; Indels 0; Gaps 0;  
Matches 733; Conservative 0; Mismatches 2;  
  
QY 1 EVKQENRLNSESSESSQGLLGYFFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLGYFFSDLNFAQPMVVTSSITGDLSPSSSELENIPSENQYF 60  
  
QY 61 QSAIWSGFIKVKSDSEYTFATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYTFATSDADNHVMTWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
  
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVTSNDLQPELKQKSSNSRKRSTASGTPVDRDN 180  
  
QY 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKKGITKYSPEKSTASDPYDFEKT 240  
Db 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKKGITKYSPEKSTASDPYDFEKT 240  
  
QY 241 GRIDKNVSPARHPLVAAPYVHVDMENILSKNEQSTQNTDSETRTSKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAPYVHVDMENILSKNEQSTQNTDSETRTSKNTSTSRHT 300  
  
QY 301 SEVHGNAEVHAFDIOGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAFDIOGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 49.1208 Seconds  
(without alignments)  
4227.791 Million cell upd

Title: US-09-848-909A-11  
 Perfect score: 3774  
 Sequence: 1 EVKQENRLNSESSESSQGL.....TSTNGIKLILFSKKGYEIG 735

Scoring table: BLOSUM62

scoring cable: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs. 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database : A_Geneseq_29Jan04.*
1: Genesexp1960s.*
2: Genesexp1990s.*
3: Genesexp2000s.*
4: Genesexp2001s.*
5: Genesexp2002s.*
6: Genesexp2003as.*
7: Genesexp2003bs.*
8: Genesexp2004s.*
```

## SIMMARTES

Result No.	Score	Query Match	Length	DB ID	Description
1	3774	100.0	735	5	AAM51494 Anthrax P
2	3761	99.7	735	5	AAM51493 Anthrax P
3	3755	99.5	735	5	AAM51485 Anthrax P
4	3754	99.5	735	5	AAM51491 Anthrax P
5	3753	99.4	735	5	AAM51495 Anthrax P
6	3750	99.4	735	5	AAM51490 Anthrax P
7	3749	99.3	735	5	AAM51487 Anthrax P
8	3749	99.3	735	5	AAM51489 Anthrax P
9	3748	99.3	735	2	AAR60179 Protective
10	3748	99.3	735	5	AAM51483 Anthrax P
11	3748	99.3	735	5	AAM51488 Anthrax P
12	3748	99.3	736	3	AAY56959 B. anthra
13	3748	99.3	763	3	AAY56960 B. anthra
14	3748	99.3	764	3	AAY56958 B. anthra
15	3748	99.3	764	4	AAB47306 Wild type
16	3748	99.3	857	7	ADE65872 Bacillus
17	3747	99.3	735	5	AAM51484 Anthrax P
18	3746	99.3	735	5	AAM51486 Anthrax P
19	3745	99.2	735	5	AAE18289 Bacillus
20	3745	99.2	764	6	AAE35717 Bacillus
21	3740	99.1	735	5	AAM51492 Anthrax P
22	3739	99.1	735	5	AAM51500 Anthrax P
23	3739	99.1	735	5	AAM51499 Anthrax P
24	3733	98.9	764	6	ABP71693 B. anthra
25	3703.5	98.9	903	2	AAR60183 PA(1-725)

Tue May 4 14:31:02 2004

Db 1395 NKNHNFINSYFDLNENEKKKINIF 1419

Search completed: May 3, 2004, 19:39:59  
Job time : 37.6261 secs







DR EMBL; AB014841; AAN35955.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 3488 AA; 411913 MW; 506F7D62999BA7B1 CRC64;  
  
Query Match 5.7%; Score 216.5; DB 5; Length 3468;  
Best Local Similarity 19.0%; Pred. No. 0.15;  
Matches 159; Conservative 130; Mismatches 275; Indels 271; Gaps 37;  
  
QY 55 SENQYFQSAIWSGFIKVKSDYVTSATSDNHTVWVDDQEV---INKASNSKIRLEK 111  
DB 160 SENEY-----TKWDDKINNSTNSG-----NKEIEFLKNTAKHSLKIKEN 202  
  
QY 112 RLYQIK-----IQYQRENTEKGLDPK---LYW---TDSQN-----KKEVISSDL 151  
DB 203 INLLKCDSDNDNLEFKQNVHLNNDVRSDFHDKNRVNNQNLKDIKDYIHDNI 262  
  
QY 152 QJPELKQKSNRKRKSTAGTVPDRNDGIPDSLEVEGIVDV-KNKRTEFLSWINI 210  
DB 263 MI-----ESNNDRNDKNCFKNTNPNYERHIIIVDTKNNGNNNNNSNPFDS--ISNI 315  
  
QY 211 HEKKGLTKYKSPKWKSTASDP-----YSDFE-KVTGRIDKNVSPPEARHPLVAAPIVHV- 264  
DB 316 NKK-----ISYPINMYNSSEDKTLNQLNLDWSILSSDSLKNAYSFSL 360  
  
QY 265 ---DMENIILSKNEDOSTQTD--SETRISKNTSTRTHTSEVHGNAEVHA--SFFDIG 317  
DB 361 EKQNSNIGKSHKNNRYRVIDEDINISKPSNISNLHDSMDNNEFNLSLCSF--- 416  
  
QY 318 GSVSAGFSNSSTVAIDSHSLAGERTWAEIWL-----NTADTARLANIRY----- 366  
DB 417 -----NSSKCVKDEMITQVGNKYKVTMDVANDNMKNSIEQHSINNESHFKNQ 468  
  
QY 367 -VNTGTAPIYVLPVTSVLGKQ-----TLATIKADE--NQLSOLAPN----- 409  
DB 469 KINSKEDANDNLSHILNNKKG:GQVNSLSDYLSIKKQESNNLSNNEALNNIV 528  
  
QY 410 -----YYPKSNLAPI-----ALNAQDFSTP--- 431  
DB 529 TNNNGSNNKNSNVYKTSQYVYFNENDINNMMHQLNLSYMKNSNNLNTSNGFKIPKN 588  
  
QY 432 -----ITWYNQFLEKTKQLRLTDQVYGNIAFYNFENGRVVDTSNWSVLPQIQ 485  
DB 589 KNIISNIDFNFIKSYKENVVVKQESNINHVEKNYTNDEINIKN---NNIENNTQ 644  
  
QY 486 ETTAR-----IIFNGKDLNLVERIAAANPSDPLETTK----- 518  
DB 645 NTTCNFINTDIIINKKKIKEYKKIDSLISLNDLSLKLBSLSSINDRYTKNNYBEK 704  
  
QY 519 --PDMTLKEAL-----KTAFGFNPNGQLQYQGHDIETFDNFQOTS 559  
DB 705 FLDDVILDDISFATSNELLQHSNYTTNHFIDNNNNNNNNINGKEDLPQNDYN---KET 761  
  
QY 560 QNIKQNLAEIATNIYTVLDKILNAKMWILIRDFHYDRNNIAVGADESUVKEAHREV 619  
DB 762 YNNIMUSENNAELF-----KISYSCDLVLGKNEIILD--RVN 800  
  
QY 620 INSTEGLLLNIDKIRKILSGVIEIETDEGLKE-----VINDRYDMLNISS--L 668  
DB 801 ENSKTE-----QVEEYQNKEDIILKLVKDDNIIIDNNYNDVNIKKDCHL 844  
  
QY 669 RODGKTIDPKKND---KLPL-----YISNPNYKN-VYAVTKENTLIINSEN 713  
DB 845 KMDNQDKINKSGQSKKKNPINNECNVHNKKNFKINEIDHPKEENTLNYESKN 899

GN UU482.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Serovar 3;  
RX MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RA Cassell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum";  
RL Nature 407:757-762(2000).  
DR EMBL; AE002145; AAF30894.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 4688 AA; 534880 MW; B53ABFAFFEE1997E CRC64;  
  
Query Match 5.5%; Score 206; DB 16; Length 4688;  
Best Local Similarity 21.3%; Pred. No. 0.68;  
Matches 174; Conservative 132; Mismatches 324; Indels 186; Gaps 40;  
  
QY 19 LLGYFSDNLNFQAPMVVTSSTGDLSPSELENIENQYFQSAIWSGFIKVKKDEYV 78  
DB 3699 LVDVYILD-NIHQNIDETRIKFDHNV-SKEIENPGVTMISKHGKWSPTDTTANFEK 3756  
  
QY 79 FATSADNHTVWVDDQEVINKASNSKIRLEKGLYQIKIYOORENTEKGLDFKLYWTD 138  
DB 3757 IETQ-----DDNDVLNIDATVKFDHNNIKQIVRIKEN-----ND 3795  
  
QY 139 SONKKEVISSDNQLPELKQKSN-----SRKKGSTAGTVPDRD----- 179  
DB 3796 WLKGGI---DNLN-PETKYKLENIELSPLKTHNTLSVINDKENISLITETGNPVLKV 3851  
  
QY 180 ---NDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTPSYD 235  
DB 3852 IQTQNDTINTQQTINTVTLGVSNSK-YNGRQIKVYADNNNVIYES---LITLQKND 3907  
  
QY 236 FEKVTGRIDKNVSPEARHPLVAAPYIVHVDMENIILSKN-EDQSTQNTQSET-RTISKNT 293  
DB 3908 YQLLSNLNEN-----REYRPEKIBINHISNTNNFEDLEKLGVSNTETITQNT 3957  
  
QY 294 STSTRHTS-EVHGNAEVHASF-----PDIGGSYSAGFS-----SNS--STVAD 335  
DB 3958 TVQWNSDSATVGRGVNFNFKISDEKILENNQQVAVFAPKETIRDTNTWLYQTRPLK 4017  
  
QY 336 HSLSLAGERTWAEIWLNT---ADTARLANIRYVNTGTAPIYVLPVTSVLGKQQLA 392  
DB 4018 DVTDFKEGIWAHDLNSNVNFKSETTYKVKIQVKNKPKAKNNINSENNVILDNTSI 4077  
  
QY 393 -----TIKADENOLSOILAPNNYYPKSNLAPIALN-AQDFSTPITMY--NQFLELE 443  
DB 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQNTINFTLGVKSKSVGKKIKLSYKNDTSEI 4137  
  
QY 444 KTKQLRLTDQVYGNIAFYNFENGR-----VRVDTGSNWSVLPQIQETTABIIFNGKD 497  
DB 4138 HINEVLESNTQTNILLNKLKNRTYTLIDVKLIDNNVSDPFKEGNLNTSITTTSA 4197  
  
QY 498 LNLVERIAAANPSDPLETT-----KPDVTLKEA-----LKIAGFNPNGNLQ 541  
DB 4198 INVNLIEISNRASTNLKSTIIKINLNDPONVLRDKDQATIVYGNKQAMGFIVSGNIK 4257  
  
QY 542 YQGKDIETFDNFDQOTSQNIK-NQLAELNATNIYTVLDKILNAKMWILIRDKRPHVD- 599  
DB 4258 YLTATVLDLNFN-DKVNIVNISFNKPSIAEN-----IGDKSNII-----YND 4304  
  
QY 600 -----RNIIAVGA---DESIVKEAHREVINSSTEGLLNIDKIRKILSGVIEIETDE 650  
DB 4305 IPKLEINNDIIVNGFINKIEIVKNAQK--NNIDVDLGLQINPKIAHNLN-FLAKFKSTN 4361  
  
QY 651 GLKEVIND--RYDMLNISSL-RODGKTFIDP-----KKY----- 681  
DB 4362 -----NDIETNVINGSSLVNNDGKYSIRFTLNNLKANKLYSLVDVYLVNNSNTIVE 4415

Db 1 MESLGINNIYNALDKLNKAKNVLVRDP-YHYDNGNIVGVDDSYLKNAYKQILNWSDD 59

Qy 626 GLLNIDKDIRKILSGYVEIDTE-----GLKEVINDRYDMLNLSLRDQ 672

Db 60 GVSLLNDEVDNALSGYMLQIKPNSHLTNSPVTITLAGKDSGVGELYRVLS-----DG 113

Qy 673 KTFIDPKYNDKPLPLYSNPYKVNVAVTKEN-TIINPSENGDPTSTNGIKILIFSKG 731

Db 114 TGLDFNKFEDENRSLV-DPGDDVYVAVTKEDFNAVTRDENGNTA-NKLKNTLVLSKI 171

Qy 732 YBI 734

Db 172 KEI 174

RESULT 10

QSRGK2

AC QSRGK2 PRELIMINARY; PRT; 1881 AA.

DT 01-JUN-2002 (T-EMBLrel. 21, Created)

DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Hemolysin.

GN FN0291.

OS Fusobacterium nucleatum (subsp. nucleatum).

CC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;

OC Fusobacterium.

OX NCBI\_TaxID=76856;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=ATCC 25586;

RC MEDLINE=21886394; PubMed=11889109;

RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,

RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,

RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,

RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,

RA Fonstein W., Kyripides N., Overbeek R.,

RT "Genome sequence and analysis of the oral bacterium Fusobacterium

RT nucleatum strain ATCC 25586."

RL J. Bacteriol. 184:2005-2018(2002).

DR EMBL; AE010541; AAL94497.1; .

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR008638; Haemagg.act.

DR InterPro; IPR005058; Peptidase\_S26.

DR InterPro; IPR001452; SH3.

DR Pfam; PF00018; SH3; 1.

DR Pfam; PF05860; Haemagg.act; 1.

DR PROSITE; PS00761; SPASE\_I\_3; 1.

KW Complete proteome.

SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;

Query Match 5.8%; Score 220.5; DB 16; Length 1881;

Best Local Similarity 20.8%; Pred. No. 0.043;

Matches 175; Conservative 131; Mismatches 303; Indels 233; Gaps 37;

Qy 30 QAPNWTST-----STTG-----DLSIPSSLE---NIPSENQY-FQSAIWSGFI 69

Db 246 QAPMYATKGVIVSSKGVYKDTQAKRDKISSTETETIGSKLAEANIKSGKTSNG 305

Qy 70 KVKKSDYTPATSDADNHTVWVDDQEVIN-----KASNSNKIRLEKGR 112

Db 306 QIRANNITINGVDSNLIPTNKDITISGNLKNSGVSSNLSNKNVKEIENSNKVVEE-K 364

Qy 113 LYQIKI-----QYQRENPTKGLDFKLYWT--DSQNKKEVISDNILQELPKQKS 161

Db 365 LSSTKIITNLGNLSAKEIKTNIFNSGKLFKNITAKDFKNGBV-SSENITTTNLENSNK 423

Qy 162 -----NSRKRSTSGAGTVPDRDNDG----- 182

Db 424 INVKENINSIVNKTNAETISKNLTNTNLDNRGNTIINNVSSGVIANNGKLLVGNTIN 483

Qy 183 ---IPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYSDFKV 239

Db 484 SQNLWTATVOGKTLDKKNKINSSGKILSDNLTDKDIFSSGNISAKVITTOELINSGEII 543

Qy 240 TGRIDKNVSPKAPRHPVAAVPIVHVDMEN-----IILSKNEDQSTQNTDSETRTSKVT 293

Db 544 SNNLSNNINNSKNIIFVGNLKIISNNLNSGVIEGLELNTSIENTGNITIKNKUTSQNL 603

Qy 294 STSRTHTSEVHGNAEVAHGFDDIGGSVSAFNSNSSTVA---IDHSLSLAGERTWAETM 350

Db 604 NNKK-----NTANVNAQFLDVHNKISS-VGNKAITMKTNNLNSGNIL-----IN 648

Qy 351 GLANTAD-----TARLNANIRYVGTAPIYVNVLTPTTSLVGLKNOTLATIKADENQLSQ 403

Db 649 SLATTAENINKSITAKNISQNVLNSGS-----VISDNITVA-----DN 687

Qy 404 ILAPNNVYPSKNLAPIALNAQKDFSTPTITMYNOFLELEKTKQLRLDLDQVYGNLATYN 463

Db 688 ITNTNNTF-----ANEKISADKIS-NSNK-LVAKNTEITKLTND---GNIVVK- 730

Qy 464 FENGREVDVTGSNNSEV-----LPQIQTETARIIFNGKOLNVLVERRIAAVNPSPDPLE 515

Db 731 -ENLKAKDITNSNTIKVGENLNTDKLQNSKTLIAKINIEKSLNNINGKITSLNANINTS 789

Qy 516 TTRPDWTLKALKIATGFNPNGLQYQKDIETEFDFNPQDQTSQNIKNQLAELNATNIY 575

Db 790 DIKNNGIIQAIK-----NINIKTSNDLKLQKGTANDSLNINAKSLENNGY--- 836

Qy 576 TVLD---KIKLNKAKNLTIRDKRFHYDRNNIAV-----GADESIVVKEAREHVINS--T 624

Db 837 --LENDGKIKFNLGTNLNNKISSSSNLTITANEISNNGVNSIIGSEANLTITANSKLN 894

Qy 625 EGLLL-----NI-DKDRIKILSGYVEIETEGLEKEVINDRYDMLNLSLRQ 670

Db 895 EGNLLFGEGIEHKLTKGTNTGTGVISSLGKLEAKVNDKHIIISDNLDTIDVNSITN 954

Qy 671 DGKTF-----IDPKK--YNDKLPYISNPYKVNVAVTKENTIIINPSENGDPTSTNGIK 722

Db 955 KGLIYSTNNKVKDFKFNLDKAEIYSSG-----DITIN-SENG-TFTNRVG 999

Qy 723 KI 724

Db 1000 DI 1001

RESULT 11

Q8II104

ID Q8II104 PRELIMINARY; PRT; 3468 AA.

AC Q8II104;

DT 01-MAR-2003 (TREMBlrel. 23, Created)

DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Hypothetical protein.

GN Pf11\_0371.

OC Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI\_TaxID=36329;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=3D7;

RC MEDLINE=22255705; PubMed=12368864;

RX Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,

RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,

RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,

RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,

RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,

RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,

RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,

RA Venter J.C., Carucci D.J., Hofman S.L., Newbold C., Davis R.W.,

RA Fraser C.M., Barrell B.;

RT "Genome sequence of the human malaria parasite Plasmodium

RT falciparum";

RL Nature 419:498-511 (2002).

```

QY 226 WSTASDPYDFKVTGRIDKNVSPPEARHPLVNAAYPIVHVDMENILSKNEQDSTQNTDSE 285
DB 231 PCTANDPYDFKVSQIDPSPVSMVARDPMISAYPIGVQMERLVVSKSE-----TITGDS 286
QY 286 TTKISQNTSTSTHSEVHGNAEVAHSPFDIGG-----SVSAGFSNSNSTVAIDHSL 338
DB 287 TKSMKSSTSHSNTNTV--GAEVGSLQLAGGIFPVFMSASANSHTQNTSVDDT- 343
QY 339 SLAGERTWAETMGLNTADTARLANAIRYNTGTAPIYVLPFTTSLVLGKQNTLATIKAE 398
DB 344 --TGE-SFSGSLGINTGESAYINPIRYNTGTAPVYVPTTTTIVDK-QSVATIKQGE 399
QY 399 NOLSQLAPNNYVPSKNLAPIALNAQKOPSSPTITMNYNQFLELEKTKQLRLDQVYGN 458
DB 400 SLIGDYLNGGTYPIIGEPFPAINTMDQFSSRLIPINYNQLKSIDNGGTVMLSSTQFTGN 459
QY 459 IATYFENGVRVDTGNSWSEVLPOIQTETARII--FNGKOLNVERIAAVNPSDPLET 516
DB 460 FAKYN-SGNLVT-DGNNWGPYLGITKSTASLTLSFGQTTQVA--VWAPNPSDPEDK 514
QY 517 TKPDMTLKALKIATFG--FNEPNGNLOVQGDIT--BEDFNEPQQTQSONIKQLAELNATN 573
DB 515 T-PKLTLEQLVKAFALEKNGKFPHGLEISKNEKIQVFLDSNTNDFENQLKNTAKD 573
QY 574 IYVLDKIKLNAQGNILIRDKRPHYDRNNIAV 605
DB 574 IMHCI--IKRN--MNLVKVITFKENISSINI 601

RESULT 8
ID Q844J8 PRELIMINARY; PRT; 775 AA.
AC Q844J8;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Vliplac.
GN Vliplac.
OS Bacillus thuringiensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1428;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi Y., Chen J., Pang Y.;
RT "Cloning of vipA(C) and vip2A(C) from Bacillus thuringiensis HB201.";
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY245547; AAC06514.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxins.
DR Pfam; PF03495; Binary toxin.
DR PRINTS; PR01391; BINARYTOXIN.
SQ SEQUENCE 775 AA; 87110 MW; 416B25394361B731 CRC64;

Query Match 19.3%; Score 728; DB 2; Length 775;
Best Local Similarity 29.1%; Pred. No. 1.3e-26;
Matches 224; Conservative 133; Mismatches 293; Indels 120; Gaps 30;

QY 4 QENRLNESSSCGLLGYFSDINTQAPMVVTSSTGDLSPSSLEN--IPSENQYFQ 61
DB 42 QKQK--QKEMDRKGLLGYFSGKDF-SNLTMPFAPTRNTLIYDQQTANKLLDKKQQYQ 97
QY 62 SATWSGPIKVKKDEYFATFADNHVTMWYDDDEVINKASNSKIRLEKGRLYQIKIQY 121
DB 98 SIRWIGLIQKKGDFTNLSEBQALIBDGIISNKGKQGVHLEKLVPIKIEYQ 157
QY 122 RENPTEKGLD-----FKLVWTSQNKKEVISSDNLQLPELKQSN----- 162
DB 158 SD--TKFNIDSKTPKFKLEFKIDSQNOQVKRDELRNPFENFKESREFLAKASKTNPFM 215
QY 163 SRKKRSTAGPTVDRDNDGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSS 222

```

```

DB 216 QKMKRDIDED---TDTGDSIPDLWEENGYT--IQNK--VAVKW-DDKFAQQGVYKYLSS 267
QY 223 PEKWSSTASDPYDFKVTGRIDKNVSPPEARHPLVNAAYPIVHVDMENILSKNEQDSTQNT 282
DB 266 PYQARTVGDPYTDWEKAAGDIPKSNAAATRNPLVAAPFSPINVDMRKILLKSDSNLS---- 323
QY 283 DSETRTISKNSTSRTHTEVHGNAEVAHSPFDIGGSVAGFSNSNS--TVAIDHSL 339
DB 324 -----NSAEAHNSNSYTYANSE-----GASIEAGFGPKGFSFGVSANYQHTET 366
QY 340 LAGERTW-----ARTMGLNTADTARLANAIRYNTGTAPIYVLPFTTSLVLGKQNTLATIK 395
DB 367 VGSQ--WGNKSKNSTEQFNSASAGYLNANYHNVGTTGGIYDAQPTTSFIL-QDSTIATIT 423
QY 396 ADENQSQLAPNNYVPSKNLAPIALNAQKDFSSPTITMNYNQFLELEKTKQLRLDQV 455
DB 424 AKSNATALSIPSGDRYPASK-EGISLTKTMDDFNSHPITLANKPQLDVAVLANNEVIKINTDQ 482
QY 456 YGNATYFENGVRVDTGNSWSEVLPOIQTETARII--FNGKOLNVERIAAVNPSDPLE 515
DB 483 DGRYGIIGVDG--KAEIGDRNSPIIDEIKGTASIIIDPADGKALETRIAAKDYKNPED 539
QY 516 TTKPDMTLKALKIATFG--FNEPNGNLOVQGDIT--ITEFDNF-----FDQQTSON 561
DB 540 KT-PSLTIKEGLIKIAYPESISEDKGLFVEYKNDGKVTKKQLSEENIMPLYLDETSKE 598
QY 562 IKQLAELNATNIYVLDKIKLNAQGNILIRDKRPHYDRNNIAVGADESIVVKEAHEVIN 621
DB 599 FERQLSDGSAGLY---DIKLTPEKNITIR-----LATVTLGFDQDSAYPWNENATW 647
QY 622 SSTEGLL-----LNIDKDIRK-----ILSGYIVE--IEDTEGLKEVI---NDR 659
DB 648 SDRFGNLRGLSLAIPQESKTYIPKDKVKNPYDYLITGYIKHFTTDSNLSGIVAFKKDN 707
QY 660 YDMLNTSSLRQDGKTFIDFKYNDKLPYISNPNYKNVYAVVTKENTIIN 709
DB 708 FEMXNMGTSIFPSONSGEGFKFTIKQ-NISG-DYILDSIQLMKRNDVN 755

RESULT 9
Q8KYK2
ID Q8KYK2 PRELIMINARY; PRT; 225 AA.
AC Q8KYK2;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Protective antigen-related protein, (pXOI-111).
GN BXA0163.
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A2012;
RX MEDLINE=22061436; PubMed=12004073;
RA Read T.D.; Salzberg S.B.; Pop M.; Shumway M.; Umayam L.; Jiang L.;
RA Holtzapple E.; Busch J.D.; Smith K.L.; Schupp J.M.; Solomon D.;
RA Keim P.; Fraser C.M.;
RT "Comparative Genome Sequencing for Discovery of Novel Polymorphisms in
RT Bacillus anthracis."; (2002).
RL Science 296:2028-2033(2002).
DR EMBL; AB011190; AAM26108.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 225 AA; 25402 MW; 2E121BE54295F9C8 CRC64;

Query Match 7.9%; Score 296.5; DB 2; Length 225;
Best Local Similarity 37.2%; Pred. No. 6.9e-07;
Matches 68; Conservative 38; Mismatches 54; Indels 23; Gaps 6;

QY 566 LAELNATNIYVLDKIKLNAQGNILIRDKRPHYDRNNIAVGADESIVVKEAHEVINSTE 625

```



```
Db 321 -----DQKTVSRNTNSKTANTAGVAINIAYQNF-----TGSITTNVSHHTENSTAVQNS 373
QY 338 LSLAGERTWAETMGNTADTARNANIRYVNTGTAPIYVNLPTTSLVLRKQNTLATIKAD 397
Db 374 ---NGB-SWNTSLKINGSAYINAVRYNTGTAPMYKVTPTNLVL-DGDTLTITKAQ 428
QY 398 ENQLSQILAPNYYSKLAPALNAQKDFSPSTPTMYNQFLEKTKQLRLDTDOYVG 457
Db 429 DNOIGNLSPNTPYKKGSLPLALNTMOQFSRLPIYNDQKLKLDAGKQIKLSTQVSG 488
QY 458 NIATYFNGRVRVDT-GSNWEVLVPOIQTETARIIFN-GKDLNLVERRIAAVNPSPLE 515
Db 489 N---YKIKNSQOIIITEGNSWSDYISQIDSLGASIIILDTGSD--VFERRVAKDSSNPED 543
QY 516 TTKPDMTLKEALKIAPGNEPNGNLQYQKDKITE--FDNFDOQTSQNIKNQLAELNATN 573
Db 544 KT-PVLTIGEATEKAFGATKNGEILYFNGMPIDESCVELIFDGTANLIKERLALNDKK 602
QY 574 IYTVLDKIKLNAKMLIIRD-----KRPY----- 598
Db 603 IYVNV---QLERGKMLIKTSTYFNFDGYNFPPSSWNSVDNNDGGLQNAANKLSGETK 658
QY 594 -----KRPY----- 598
Db 659 IIVPMKLNPKRYVFPFSGYLKNSSTSPNITVNIKABQKTYNLVSENDYKFSYEFETIG 718
QY 599 -DRNN-----IAGADESWKKEAHRV-----I 620
Db 719 RDSNIEIILTSSGTIFDLNLSITELNSTPEILKSPDIKVPDQELI-DAHKYVADLSF 777
QY 621 NSSTEGLLN-----IDKDIRKILSGYIVIE-DTEGLKEVINDRYMDMLNISLRQD 671
Db 778 NQSTANVYLDGLYFEPTQNTKEVLDYIQYKVEATLEYSGFKDICTGDKELANYTGDSNQ 837
QY 672 GKT-FIDEKYNKDLPLYISNPNY-----KVVVAVTKENTIINPSENGDTSTNGIKKILI 726
Db 838 PKTYNVNFRSY-----FISGENVMYKRLRIYAITPEN-----KELLV 875
QY 727 FS 728
Db 876 LS 877

RESULT 5
Q9KH41 ID Q9KH41 PRELIMINARY; PRT; 876 AA.
AC Q9KH41;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CdcB.
GN CDTB.
OS Clostridium difficile.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1496;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 20309;
RA Chang S.Y., Song K.P.;
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF271719; AAF81761.1; -.
DR HSSP; P13423; 1ACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;
```

```
Query Match 23.3%; Score 880.5; DB 2; Length 876;
Best Local Similarity 31.6%; Pred. No. 9e-34;
Matches 251; Conservative 139; Mismatches 280; Indels 125; Gaps 30;

QY 4 QENRLINESSESSQGLGYFSDLNFOAPMVYTSSTGDSIPSSBELEN-IPSENOYFQS 62
Db 40 KKKEIVNEIDILPNGLMGYYFTDEHFQDLKMAPIKDGNLKEPKKVDKLLDKSDVKS 99
QY 63 AHWGIGIKVKKDEYTFATSAADNHVTMWYDDQEVINKASNSKIRLEKGRVCIQIOYOR 122
Db 100 IRMTGILIPSKOGGYFLSDRDD-VLMQVNTSTST---SNTLKNVKKGKFKVRIELO 155
QY 123 EN--PTEKGLDFKLYWTDQSNKKEVISSDNLOLPELKQKSSNRKRSSTAGTVP---- 176
Db 156 KNLGSIDNLSPLXW-ELDGMKKIIPENLFLRDY-----SNIKD-----DPFIPNNNF 205
QY 177 -----DRDNGIDPSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSS 222
Db 206 FDPKLSMDWEDELDTDNDNIPOSYERNGYTI-----KLIIVKWDSEFAE-QYKKYKVN 250
QY 223 PEKWTASDPYSDFEKTGRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTONT 282
Db 261 YLESNTAGDPTDYERKASGDFDKAIKTEARDPLVAAPIVGVGMEKLIISTNEHAST--- 317
QY 283 DSETRISKNTSTSRTHTEVHGNAEVHASFPDIGSVSAGFSNSNSTVAIDHS----- 337
Db 318 -DQKTVSRATNTSKTESNTA-----GVSVMVGQNGTANVTNYSHTTON 363
QY 338 ---LSLAGERTWAETMGNTADTARNANIRYVNTGTAPIYVNLPTTSLVLRKQNTLATI 394
Db 364 STAVQNSGSEWNTGSIKNGESAYINAVRYNTGTAPMYKVTPTNLVL-DGDTLTIT 422
QY 395 KADENQLSQILAPNYYPSKQLAPIALNAQKDFSPSTPTMYNQFLEKTKQLRLDTDO 454
Db 423 KAENQIGNNLSPGDTVPKGLSPLALNTMDQFSRLPIYNDQKLKLDAGKQIKLETQ 482
QY 455 VYGNIAIYFNGRVRVDTGNSWSEVLVPOIQTETARIIFNGKDLNLVERRIAAVNPSDPL 514
Db 483 VSGNFGTKN-SSQOI-VTEGNSWSDYISQIDSLGASIIILDTEN-ESYERRVTKNLQDPE 539
QY 515 ETTKPDMTLKEALKIAPGNEPNGNLQYQKDKITE--FDNFDOQTSQNIKNQLAELNAT 572
Db 540 DKT-PELTIGEATEKAFGATKDGILYFNDIPIDESCVELIFDDNTANKIKSLKLSDK 598
QY 573 NIYTVLDKIKLNAKMLIIRDKRPY---DRNNIAGADESVVKEAHRVINGSTEG--- 627
Db 599 KIYVNV---KLERGMNLIKITPTFTNFDDYNNYP--STWSNVNTNKGDLQGSANKLNG 652
QY 628 -----LLNIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656
Db 653 ETKIKIPMSSELKPKRYVVPFSGYKSDPLTNSIIVKIKAKEEKTDLVLPQGYTKFSYEFE 712
QY 657 NDRYDMLNIS-SLRQDKTFTIDFKYND--KLPLYISNPNYK-----VNVVAVTKENT 706
Db 713 TTEKSSNIEITLIGSGTITLDNLSTELNSTPEILDEPEVKIPTDQEIIDAHKIYFADL 772
QY 707 IINPSENGDTSTNGI 721
Db 773 NFNFS-TGNTYINGM 786

RESULT 6
Q46221 ID Q46221 PRELIMINARY; PRT; 875 AA.
AC Q46221;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Iota toxin component Ib precursor.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
CX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCUG 20309;
RA Chang S.Y., Song K.P.;
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF271719; AAF81761.1; -.
DR HSSP; P13423; 1ACC.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0015070; F:toxin activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR003896; Anthrax toxinB.
DR Pfam; PF03495; Binary toxinB.
DR PRINTS; PR01391; BINARYTOXINB.
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;
```

AC 032739;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE ADP-ribosyltransferase.  
GN CDTB.  
OS Clostridium difficile.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1496;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CD196;  
RX MEDLINE=97230316; PubMed=9119480;  
FA Perelle S., Gibert M., Bourlioux P., Corthier G., Popoff M.R.;  
RT "Production of a complete binary toxin (actin-specific ADP-  
ribosyltransferase) by Clostridium difficile CD196.";  
RL Infect. Immun. 65:1402-1407(1997).  
DR EMBL; L76081; AAB67305.1; -.  
DR HSSP; P13423; IACC.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0015070; P:toxin activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0003405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax\_toxinB.  
DR Pfam; PF03495; Binary\_toxB; 1.  
DR PRINTS; PR01391; BINARYTOXINB.  
KW Transferase.  
SQ SEQUENCE 876 AA; 98797 MW; 25E06B2D45CE2B3B CRC64;  
  
Query Match 23.4%; Score 882.5; DB 2; Length 876;  
Best Local Similarity 31.7%; Pred. No. 7.3e-34;  
Matches 252; Conservative 138; Mismatches 280; Indels 125; Gaps 30;  
  
QY 4 QENRLNESESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELEN-IPSENQYFQS 62  
DB 40 KKEIVNEILPNGLMGVYFSDHFQDLKMAPKDGKLNKPEKKVGLDKDKSDVKS 99  
  
QY 63 AIWSGFIVKKSDEYTFATSDAHNVMTWDDQEVINKASNNKIRLEKGLYQIKQYOR 122  
DB 100 IRWTGRIPSKGEYTLSTRDD-VLMQVNTSTI---SNTLKNVKKGKEYKVRLELD 155  
  
QY 123 EN--PTEKGLDFKLYWTDQNKKEVISSDNLOLPELKQSSNRKRSTSAQTVVP---- 176  
DB 156 KNLGSDNLSSPNLYW-ELDGMKKIIPENLFLRDY----SNIKD----DPFIPNNF 205  
  
QY 177 -----DRDNDGIPDSLEVEGYTVDVKNKIFLSPWISNHEKKGLKYKSS 222  
DB 206 FDPKLMSDWEDELDTDNDNIPDSYERNGYTI-----KDLIAVKWEDSFAE-QGYKYYVN 260  
  
QY 223 PEKWTASDPSPDFEKTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNT 282  
DB 261 YLESNTAGDPTDYEKASGDFKAIKTEARDPLVAAYPIVGVMEKLIISTNEHAST--- 317  
  
QY 283 DSETRISKVNTSRTHTSEVHGNAEVHASFDDIGSVSAGSFSNSSTVAIDHS----- 337  
DB 318 -DQKTVSRATNTSKTESNTA-----GVSANVQYQNGFTANVTNYSHHTDN 363  
  
QY 338 ---LSLAGETMAETWGLNTADTARLANIRVYNTGTAPIYVLTPTSLVLGKNTLATI 394  
DB 364 STAVQDSNGESWTGSLNGESAYINAVRYNTGTAPYKVTPTNLVL-DGDTLSIT 422  
  
QY 395 KADENQLSQILAFNNYPSKNLAPIALNAQKQDSSTPTITWYNNQFLEBKTQLRLDTPQ 454  
DB 423 KAQENQIGNLSPGDTVPKGLSPLALNTMDQFSRLIPINYPDLKLDAGKQIKLETQ 482  
  
QY 455 VYGNATYFENGVRVDTGNSHSEVLPOIQTETARIIFNGKDLNLVERRIAVNPSDPL 514  
DB 483 VSGNFGTKN--SSQOI-VTEGNSWSDYISQDSISASILDTEN-ESYERAVTAKNLQDPE 539  
  
QY 515 ETTKPDMLKALKIAFGFNEPNGLQYQKDKITE--FDNFQDQTSQNIKNGLAELNAT 572  
DB 540 DKT-PELTIGEATEKAFKQDGLLYFNIPIDESCVELLPDNTANKIKDLSKLTSDK 598

QY 573 NIYTVLDKIKLNKQWILIRDKRFHY---DRNNIAGADESVVKEAHEVINSSTEGL-- 627  
DB 599 KIYV-----KLERGMNLIKTYFTNFDYNNYP--STWSNVNTNQDGLQGSANKLNG 652  
  
QY 628 -----LLNIDKDIRKILSGY-----IVEI-----EDTEGL-----KVI 656  
DB 653 ETKIKIPSELKPKYKRYFSGYKDXPLTSNSIIVIKAKKEKTDLYLVPEQGYTKSYEF 712  
QY 657 NDYMDLANIS-SLRQDGKTFIDFKKYND--KLPLYISNPNYKV-----NRYAVTKNT 706  
DB 713 TTEKSSNIEITLIGSGTTLNLSITELNSTPEILDPEVKIPTDQDQIMDAHKYFADL 772  
  
QY 707 IINPSENGDTSTNGI 721  
DB 773 NFNPS-TGNTYINGM 786  
  
RESULT 4  
ID 006498 PRELIMINARY; PRT; 879 AA.  
AC 006498;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Sb component.  
GN SSS.  
OS Clostridium spiroforme.  
OC Bacteria; Firmicutes; Mollicutes.  
OX NCBI\_TaxID=29348;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CS246;  
RA Gibert M., Berelle S., Daube G., Popoff M.R.;  
RT "Clostridium spiroforme toxin genes are related to C. perfringens iota  
toxin genes but have a different genomic localization.";  
RL Syst. Appl. Microbiol. 20:337-347(1997).  
DR EMBL; X97969; CAA66612.1; -.  
DR HSSP; P13423; IACC.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0015070; P:toxin activity; IEA.  
DR GO; GO:0003405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax\_toxinB.  
DR Pfam; PF03495; Binary\_toxB; 1.  
DR PRINTS; PR01391; BINARYTOXINB.  
SQ SEQUENCE 879 AA; 98739 MW; 40685ACB8E05BA01 CRC64;  
  
Query Match 23.4%; Score 882.5; DB 2; Length 879;  
Best Local Similarity 28.8%; Pred. No. 7.3e-34;  
Matches 260; Conservative 146; Mismatches 259; Indels 237; Gaps 34;  
  
QY 4 QENRLNESESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQ-YFQS 62  
DB 36 QEVETTNEKTVSSNGLMGVYFADEHFKDLLEMAPVKNGELKFEKNKVEKLTTEKTNKS 95  
  
QY 63 AIWSGFIVKKSDEYTFATSDAHNVMTWDDQEVINKASNNKIRLEKGLYQIKQYOR 122  
DB 96 IRWTGRIPSKGEYTLSTRDD-VLMQVNTSTI---SNTLKNVKKGKEYKVRLELD 150  
  
QY 123 ENPTEKGLDF-----KLYWTDQNKKEVISSDNLOLPELKQSSN----- 162  
DB 151 ----DKOIGYVDDLSLSPKLYW-BLNGDKTLIPEKNLFLRDYSKIDENDPFIPKDNFPLK 205  
  
QY 163 ---SKKQSTSA-GTVDPDRNDGIPDSLEVEGYTVDVKNKETFSLSPISNHEKKGLTKY 219  
DB 206 LKRSARLASAGWGEDLDTDNDNIPDAYEKNGYTI-----KDSIAVKW-EDSFAQCYK 260  
  
QY 220 KSSPEKNTASDPYSDFEKTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQST 279  
DB 261 LSSVLESTAGDPTDYKASGDFKAIKTEARDPLVAAYPIVGVMEKLIISTNEHAST 320  
  
QY 280 QNTDSETRISKVNTSRTHTSEVHGNAEVHASFDDIGSVSAGSFSNSSTVAIDHS 337

Db	20	EVKQENRLNSESSESSQGLLYGYSFSLDNFQAPMVVTSSTTGDLSIPSELENIPSENQYF	79
Qy	61	QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY	120
Db	80	QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY	139
Qy	121	QRENPTKGLDFKLYWTDSONKKEVISSDNQLPDLKQKSSNSRKKSTASGPTVPDRN	180
Db	140	QRENPTKGLDFKLYWTDSONKKEVISSDNQLPDLKQKSSNSRKKSTASGPTVPDRN	199
Qy	181	DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPYDFEKT	240
Db	200	DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPYDFEKT	259
Qy	241	GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT	300
Db	260	GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT	319
Qy	301	SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	320	SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	379
Qy	361	NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILLAPNNYPSKNLAPIA	420
Db	380	NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILLAPNNYPSKNLAPIA	439
Qy	421	LNAQKDFSSFTPTIMYVNTQFLEKTKQLRLDQVYGNIAVYVNTKNTIINPSENGDTSTNG	480
Db	440	LNAQKDFSSFTPTIMYVNTQFLEKTKQLRLDQVYGNIAVYVNTKNTIINPSENGDTSTNG	499
Qy	481	LPOIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLKEALKIATGFNEPKNL	540
Db	500	LPOIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLKEALKIATGFNEPKNL	559
Qy	541	OYQKDIITEFDNFDOOTSONIKNQLAELNATVIYVLDKIKLNAKMNILIRDKRFHYDR	600
Db	560	OYQKDIITEFDNFDOOTSONIKNQLAELNATVIYVLDKIKLNAKMNILIRDKRFHYDR	619
Qy	601	NNIAGADESVVKEAHREVNINSTEGLLLINIDKDIRKILSGYVIEIDTEGLKEVINDRY	660
Db	620	NNIAGADESVVKEAHREVNINSTEGLLLINIDKDIRKILSGYVIEIDTEGLKEVINDRY	679
Qy	661	DLNLISSLRQDGKTFIDFKYNDKPLIYISNPYKVVVAVTKNTIINPSENGDTSTNG	720
Db	680	DLNLISSLRQDGKTFIDFKYNDKPLIYISNPYKVVVAVTKNTIINPSENGDTSTNG	739
Qy	721	IK 722	
Db	740	IK 741	
RESULT 2			
Q937W3	ID	Q937W3	PRELIMINARY; PRT; 743 AA.
AC	Q937W3;	AT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Pag protein (Fragment).	GN	PAG.
OS	Bacillus anthracis.	OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX	NCBI_TaxID=1392;	RN	(1)
RC	SEQUENCE FROM N.A.	RP	STRAIN=Carbosap.
RA	Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,	RA	Fasanella A., Francia M., Ciuchini F.;
RT	"Sequence analysis of the genes encoding for the major virulence	RT	factors of Bacillus anthracis vaccine strain 'Carbosap.';"
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.		

DR	EMBL; AJ413936; CAC93934.1; ..	
DR	GO; GO:0005576; C:extracellular; IEA.	
DR	GO; GO:0046821; C:extrachromosomal DNA; IEA.	
DR	GO; GO:0015070; P:toxin activity; IEA.	
DR	GO; GO:0009405; P:pathogenesis; IEA.	
DR	InerPro; IPR003896; Anthrax toxinB.	
DR	Pfam; PF03495; Binary_coxB; I.	
DR	PRINTS; PR01391; BINARYTOXINB.	
KW	Plasmid.	
FT	NON TER 1 743	
FT	NON TER 743 743	
SEQ	SEQUENCE 743 AA; 83363 MW; 292757AD2D5DE5A6 CRC64;	
Query Match		
Best Local Similarity 97.7%; Score 3689; DB 2; Length 743;		
Matches 718; Conservative 1; Mismatches 3; Indels 0; Gaps 0;		
Qy	1	EVKQENRLNSESSESSQGLLYGYFDLPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db	22	EVKQENRLNSESSESSQGLLYGYFDLPQAPMVVTSSTTGDLSIPSELENIPSENQYF 81
Qy	61	QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db	82	QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 141
Qy	121	QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKKRSTASGPTVPDRN 180
Db	142	QRENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNSRKKRSTASGPTVPDRN 201
Qy	181	DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPYDFEKT 240
Db	202	DGIPDSLEVEGYTVDVKNKRTPLSPWISNIHEKGLTKYKSSPEKWSSTASDPYDFEKT 261
Qy	241	GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db	262	GRIDKNVSPKARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 321
Qy	301	SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db	322	SEVHGNAEVSASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 381
Qy	361	NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
Db	382	NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILLAPNNYPSKNLAPIA 441
Qy	421	LNAQKDFSSFTPTIMYVNTQFLEKTKQLRLDQVYGNIAVYVNTKNTIINPSENGDTSTNG 480
Db	442	LNAQKDFSSFTPTIMYVNTQFLEKTKQLRLDQVYGNIAVYVNTKNTIINPSENGDTSTNG 501
Qy	481	LPOIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLKEALKIATGFNEPKNL 540
Db	502	LPOIQTETARIIIFNGKDLNVERRIAANVPSDPLETTPDWTLKEALKIATGFNEPKNL 561
Qy	541	OYQKDIITEFDNFDOOTSONIKNQLAELNATVIYVLDKIKLNAKMNILIRDKRFHYDR 600
Db	562	OYQKDIITEFDNFDOOTSONIKNQLAELNATVIYVLDKIKLNAKMNILIRDKRFHYDR 621
Qy	601	NNIAGADESVVKEAHREVNINSTEGLLLINIDKDIRKILSGYVIEIDTEGLKEVINDRY 660
Db	622	NNIAGADESVVKEAHREVNINSTEGLLLINIDKDIRKILSGYVIEIDTEGLKEVINDRY 681
Qy	661	DLNLISSLRQDGKTFIDFKYNDKPLIYISNPYKVVVAVTKNTIINPSENGDTSTNG 720
Db	682	DLNLISSLRQDGKTFIDFKYNDKPLIYISNPYKVVVAVTKNTIINPSENGDTSTNG 741
Qy	721	IK 722
Db	742	IK 743
RESULT 3		
O32739		
ID	O32739	PRELIMINARY; PRT; 876 AA.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds  
(without alignments)  
6697.424 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENLLNSESQQLL.....TSTNGIKLIFSKGVEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3689	97.7	741	2 Q937W2	Q937W2 bacillus an
2	3689	97.7	743	2 Q937W3	Q937W3 bacillus an
3	882.5	23.4	876	2 Q32739	Q32739 clostridium
4	882.5	23.4	879	2 Q06498	Q06498 clostridium
5	880.5	23.3	876	2 Q9KH41	Q9KH41 clostridium
6	876	23.2	875	2 Q46221	Q46221 clostridium
7	818	21.7	721	2 Q86171	Q86171 clostridium
8	728	19.3	775	2 Q844J8	*Q844J8 bacillus th
9	296.5	7.9	225	2 Q8KYK2	Q8KYK2 bacillus an
10	220.5	5.8	1881	16 Q8RGK2	Q8RGK2 fusobacteri
11	216.5	5.7	3468	5 Q8II04	Q8II04 plasmodium
12	206	5.5	4688	16 Q9PQ08	Q9PQ08 ureaplasma
13	203.5	5.4	1904	5 Q8IKS2	Q8IKS2 plasmodium
14	200.5	5.3	2849	5 Q8IHY4	Q8IHY4 plasmodium
15	197.5	5.2	2940	5 Q8IHP9	Q8IHP9 plasmodium
16	195	5.2	1404	2 Q8RJN9	Q8RJN9 mycoplasma

17	192.5	5.1	675	16	Q897H6	Q897H6 clostridium
18	192.5	5.1	1301	5	Q8WSKS	Q8WSKS plasmodium
19	190.5	5.0	2771	5	Q26216	Q26216 plasmodium
20	189.5	5.0	2444	16	Q81GX1	Q81GX1 bacillus ce
21	189	5.0	2792	5	Q814R2	Q814R2 plasmodium
22	189	5.0	9904	16	Q8NWQ6	Q8NWQ6 staphylococ
23	188	5.0	3248	5	Q8ICP9	Q8ICP9 plasmodium
24	187.5	5.0	1301	5	Q8IHQ2	Q8IHQ2 plasmodium
25	187.5	5.0	1681	5	Q8I548	Q8I548 plasmodium
26	186.5	4.9	1182	16	Q8EW90	Q8EW90 mycoplasma
27	186.5	4.9	2235	5	Q8IB09	Q8IB09 plasmodium
28	186	4.9	4261	5	Q8IFP4	Q8IFP4 plasmodium
29	185	4.9	3504	5	Q8II45	Q8II45 plasmodium
30	184.5	4.9	1125	16	Q88PM9	Q88PM9 mycoplasma
31	184.5	4.9	2529	16	Q25579	Q25579 helicobacte
32	184	4.9	1859	5	Q8IC27	Q8IC27 plasmodium
33	184	4.9	3203	5	Q7YWE7	Q7YWE7 plasmodium
34	184	4.9	3203	5	Q7YWE6	Q7YWE6 plasmodium
35	183	4.8	1936	5	Q8I5A6	Q8I5A6 plasmodium
36	183	4.8	2747	5	Q9BJX9	Q9BJX9 plasmodium
37	182.5	4.8	1387	5	Q9GZ76	Q9GZ76 plasmodium
38	182.5	4.8	3080	5	Q7YWF0	Q7YWF0 plasmodium
39	182	4.8	6713	16	Q99U54	Q99U54 staphylococ
40	182	4.8	6713	16	Q931R6	Q931R6 staphylococ
41	181	4.8	2369	5	Q26223	Q26223 plasmodium
42	181	4.8	4152	2	Q9ZHL3	Q9ZHL3 haemophilus
43	181	4.8	4152	16	Q7VLE8	Q7VLE8 haemophilus
44	181	4.8	10061	5	Q8I3Z1	Q8I3Z1 plasmodium
45	180.5	4.8	149	2	Q9RM77	Q9RM77 clostridium

#### ALIGNMENTS

#### RESULT 1

Q937W2 ID Q937W2 PRELIMINARY; PRT; 741 AA.  
AC Q937W2; 01-DEC-2001 (T-EMBLrel. 19, Created)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
DE Pag protein (Fragment).  
GN PAG.  
OS Bacillus anthracis.  
OG Plasmid pX01.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ferrara;  
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,  
RA Fasanella A., Francia M., Ciuchini F.,  
RT "Sequence analysis of the genes encoding for the major virulence  
RT factors of bacillus anthracis vaccine strain 'Carbosap'";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ413937; CAC93935.1;  
DR GO; GO:000576; C:extracellular; IEA.  
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax toxinB.  
DR Pfam; PF03495; B:Binary COXB; I.  
DR PRINTS; PR01391; BINARYTOXINE.  
KW Plasmid.  
FT NON TER 1 741  
FT NON TER 741 741  
SQ SEQUENCE 741 AA; 83153 MW; C7F95820B73065C0 CRC64;

Query Match 57.7%; Score 3689; DB 2; Length 741;

Best Local Similarity 99.4%; Pred. No. 7.8e-167;

Matches 718; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EVKQENLLNSESQQLLYGVFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

1	3773	100.0	735	5	AAM51491	Anthrax P
2	3769	99.9	735	5	AAM51490	Anthrax P
3	3768	99.9	735	5	AAM51489	Anthrax P
4	3767	99.8	735	5	AAM51488	Anthrax P
5	3767	99.8	735	5	AAR60179	Protectiv
6	3767	99.8	735	5	AAM51483	Anthrax P
7	3767	99.8	735	5	AAM51488	Anthrax P
8	3767	99.8	735	5	AAM51493	Anthrax P
9	3767	99.8	736	3	AAY56959	B. anthra
10	3767	99.8	764	3	AAY56960	B. anthra
11	3767	99.8	764	3	AAY56958	B. anthra
12	3767	99.8	857	7	AAB47306	Wild type
13	3764	99.8	735	5	ADE65872	Bacillus
14	3764	99.8	764	6	AAB35717	Bacillus
15	3763	99.7	735	5	AAM51487	Anthrax P
16	3761	99.7	735	5	AAM51484	Anthrax P
17	3761	99.7	735	5	AAM51485	Anthrax P
18	3759	99.6	735	5	AAM51486	Anthrax P
19	3759	99.6	735	5	AAM51495	Anthrax P
20	3759	99.6	735	5	AAM51492	Anthrax P
21	3758	99.6	735	5	AAM51500	Anthrax P
22	3758	99.6	735	5	AAM51499	Anthrax P
23	3754	99.5	735	5	AAM51494	Anthrax P
24	3752	99.4	764	6	ABP71693	B. anthra
25	3722.5	98.7	903	25	AAR60183	PA(1-725)

```

XX SQ Sequence 735 AA;
Query Match 100.0%; Score 3773; DB 5; Length 735;
Best Local Similarity 100.0%; Pred. No. 6.9e-245;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMWVTSSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMWVTSSTTGDLSIPSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDRYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDRYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGPTVPDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNISRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNISRTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYNTGTAPIYVNLPTSLVLGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYNTGTAPIYVNLPTSLVLGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480

QY 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDDPLETKCDMTLKEALKIARFNPENGNI 540
DB 481 LPQIQTETARIIFNGKDLNLVERRIAANVPSDDPLETKCDMTLKEALKIARFNPENGNI 540

QY 541 QYQKQDITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLAKWNLILRDKRFHYDR 600
DB 541 QYQKQDITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLAKWNLILRDKRFHYDR 600

QY 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660

QY 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNKYVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNLSLRQDGKTFIDFKYNDKPLIYISNPNKYVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILFESKXGYEIG 735
DB 721 IKKILFESKXGYEIG 735

```

## RESULT 2

```

AAM51490
ID AAM51490 standard; protein; 735 AA.
XX
AC
XX
XX
XX 01-FEB-2002 (first entry)
DE Anthrax PA mutant D425B.
XX
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.
KW
XX

```

```

OS Bacillus anthracis.
OS Synthetic.
PH Key Location/Qualifiers
FT Misc-difference 425 /note= "Wild-type Asp substituted by Glu"
XX
XX WO200182788-A2.
XX 08-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014372.
XX
XX 04-MAY-2000; 2000US-0201800P.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Collier RJ, Sellman BR;
XX
XX WPI; 2002-017725/02.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX
XX Claim 4; Page; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
XX Sequence 735 AA;
Query Match 99.9%; Score 3769; DB 5; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.3e-244;
Matches 734; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSESSQGLLYGYFSDLNFOAPMWVTSSTTGDLSIPSELENIPSENQYF 60
DB 1 EVKQENLLNESSESSQGLLYGYFSDLNFOAPMWVTSSTTGDLSIPSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDRYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDRYTATGADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGPTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGPTVPDRDN 180

QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYSDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNISRTHT 300
DB 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNISRTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYNTGTAPIYVNLPTSLVLGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYNTGTAPIYVNLPTSLVLGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480

```

Query Match 99.9%; Score 3768; DB 5; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.5e-244;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSESSQGLLYGFFSDLPQAPMVTTSITGDLSPSSSELENI PSENQYF 60  
DB 1 EVKQENRLNESSESSQGLLYGFFSDLPQAPMVTTSITGDLSPSSSELENI PSENQYF 60  
QY 61 QSAWSGFIKVKKSDBYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
DB 61 QSAWSGFIKVKKSDBYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKOKSSNSRKRSTSGPTVPDRN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKOKSSNSRKRSTSGPTVPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPSDEPKYT 240  
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPSDEPKYT 240  
QY 241 GRIDKNVSPARHPLVAAPYIHHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
DB 241 GRIDKNVSPARHPLVAAPYIHHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVLPPTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYVLPPTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGSNWEV 480  
DB 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGSNWEV 480  
QY 481 LPQIQTETARIIFNGKDLNVERRIAAVNPSPDETTKPDMLTKALKIAFGNEPENG 540  
DB 481 LPQIQTETARIIFNGKDLNVERRIAAVNPSPDETTKPDMLTKALKIAFGNEPENG 540  
QY 541 QYQKDIETEFDFNFQDQTSQNIKNOLAEALNATNYTVLDKIKLAKNKNILIRDKRPHYDR 600  
DB 541 QYQKDIETEFDFNFQDQTSQNIKNOLAEALNATNYTVLDKIKLAKNKNILIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIDKIRKILSGYIVEIDTEGLKEVINDRY 660  
DB 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIDKIRKILSGYIVEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 3  
ID AAM51489 standard; protein; 735 AA.  
XX AAM51489;  
XX 01-FEB-2002 (first entry)  
XX Anthrax PA mutant D425N.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX Bacillus anthracis.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Misc-difference 425 /note= "Wild-type Asp substituted by Asn"  
XX WO200182788-A2.  
XX 08-NOV-2001.  
XX 04-MAY-2001; 2001WO-US014372.  
XX 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX Claim 4; Page; 77pp; English.  
XX The invention relates to antibacterial agents comprising mutant forms of  
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX moiety is anthrax protective antigen (PA) and using these mutants or  
XX compositions of them for protecting against Bacillus anthracis infections  
XX in humans, especially as vaccines. Note: The present sequence is not  
XX given in the specification but is derived from the Bacillus anthracis  
XX wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX Sequence 735 AA;  
SQ

QY 1 EVKQENRLNESSESSQGLLYGFFSDLPQAPMVTTSITGDLSPSSSELENI PSENQYF 60  
DB 1 EVKQENRLNESSESSQGLLYGFFSDLPQAPMVTTSITGDLSPSSSELENI PSENQYF 60  
QY 61 QSAWSGFIKVKKSDBYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
DB 61 QSAWSGFIKVKKSDBYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKOKSSNSRKRSTSGPTVPDRN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVTSNLOLPELKOKSSNSRKRSTSGPTVPDRN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPSDEPKYT 240  
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPSDEPKYT 240  
QY 241 GRIDKNVSPARHPLVAAPYIHHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
DB 241 GRIDKNVSPARHPLVAAPYIHHVDMENILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVLPPTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYVLPPTSLVGLKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGSNWEV 480  
DB 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDTPQVYGNATYVNFENGVRVDTGSNWEV 480  
QY 481 LPQIQTETARIIFNGKDLNVERRIAAVNPSPDETTKPDMLTKALKIAFGNEPENG 540  
DB 481 LPQIQTETARIIFNGKDLNVERRIAAVNPSPDETTKPDMLTKALKIAFGNEPENG 540  
QY 541 QYQKDIETEFDFNFQDQTSQNIKNOLAEALNATNYTVLDKIKLAKNKNILIRDKRPHYDR 600  
DB 541 QYQKDIETEFDFNFQDQTSQNIKNOLAEALNATNYTVLDKIKLAKNKNILIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIDKIRKILSGYIVEIDTEGLKEVINDRY 660  
DB 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIDKIRKILSGYIVEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNSSLRQDGKTFIDFKKNDKLPYISNPYKVNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 4  
AAR60179  
ID AAR60179 standard; protein; 735 AA.  
XX AAR60179;  
XX 25-MAR-2003 (revised)  
XX 03-APR-1995 (first entry)  
XX Protective antigen of Bacillus anthracis.  
XX Anthrax; Bacillus anthracis; fusion protein; protective antigen;  
XX protective antigen; cell killing; targeting; targeting; pathogen;  
XX intracellular; HIV; human immunodeficiency virus; toxin.  
XX Bacillus anthracis.

XX W09418332-A2.  
XX PD 18-AUG-1994.  
XX PF 14-FEB-1994; 94WO-US001624.  
XX PR 12-FEB-1993; 93US-00021601.  
XX PR 25-JUN-1993; 93US-00082849.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Leppia SH, Klimpel K, Arora N, Singh Y, Nichols PJ;  
XX WPI; 1994-279753/34.  
XX DR N-PSDB; AAQ70180.  
XX  
PT Nucleic acid encoding anthrax toxin fusion protein - useful for  
PT targeting toxin to specific cells, eg for killing tumour cells or HIV-  
PT infected cells.  
XX  
XX Disclosure; Page 81-83; 124pp; English.  
XX  
CC The sequence encoding the protective antigen of Bacillus anthracis may be  
CC used in the construction of a nucleic acid which encodes a fusion protein  
CC comprising the anthrax protective antigen binding domain of the native  
CC anthrax lethal factor and a sequence encoding an activity inducing domain  
CC of a second protein. The fusion proteins are useful for the specific  
CC killing of tumour cells or the killing of cells infected with  
CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to  
CC correct FN field.)  
XX  
XX Sequence 735 AA;  
XX  
Query Match 99.8%; Score 3767; DB 2; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.7e-244;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSSSGLLGYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSSSGLLGYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDVETPATSADNHVTMMVDQEVINKASNKIRLEKGLYQIKIY 120  
DB 61 QSAIWSGFIKVKSDVETPATSADNHVTMMVDQEVINKASNKIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQSSNSRKRSTSGPTVDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLPELKQSSNSRKRSTSGPTVDRDN 180  
QY 181 DGIPTSLVEGYTVDVKNKTEFLSPWISNHEKGLTKYKSPKKNSTASDPYDPEKYT 240  
DB 181 DGIPTSLVEGYTVDVKNKTEFLSPWISNHEKGLTKYKSPKKNSTASDPYDPEKYT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFIDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVHASFIDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPYNYVLPITSLVLGKNQTLATIKAKENQLSILAPNNYPSKNLAPIA 420  
DB 361 NANIRYNTGTAPYNYVLPITSLVLGKNQTLATIKAKENQLSILAPNNYPSKNLAPIA 420  
QY 421 LNAQDSSSTPIWNNYQFLEKTKQLRLDQVYGNATYFNGRVRVDTGGSNWSEV 480  
DB 421 LNAQDSSSTPIWNNYQFLEKTKQLRLDQVYGNATYFNGRVRVDTGGSNWSEV 480  
QY 481 LPOICETTARIIFNGKOLNVERIAAANPSDPLETTKPDWTLKEALKIAFGNENGNL 540  
DB 481 LPOICETTARIIFNGKOLNVERIAAANPSDPLETTKPDWTLKEALKIAFGNENGNL 540

QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLKDIKLNAMNILLIRKRFHYDR 600  
DB 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLKDIKLNAMNILLIRKRFHYDR 600  
QY 601 NNTIAGADESVKAEHREVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660  
DB 601 NNTIAGADESVKAEHREVINSSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKKYNKDKLPLYISNPYKVNVAVTYKNTIINPSENGDSTNG 720  
DB 661 DMLNISSLRQDGKTFIDFKKYNKDKLPLYISNPYKVNVAVTYKNTIINPSENGDSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
DB 721 IKKILIFSCKGYEIG 735  
RESULT 5  
AAMS1483  
ID AAMS1483 standard; protein; 735 AA.  
XX  
XX AAMS1483;  
XX  
XX 01-FEB-2002 (first entry)  
XX Anthrax PA protein.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.  
XX Bacillus anthracis.  
XX WO200182788-A2.  
XX 08-NOV-2001.  
XX 04-MAY-2001; 2001WO-US014372.  
XX 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX DR N-PSDB; AA199904.  
XX Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX Disclosure; Fig 13; 77pp; English.  
XX  
CC The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. The present sequence is that of the  
CC anthrax PA protein  
XX  
SQ Sequence 735 AA;  
Query Match 99.8%; Score 3767; DB 5; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.7e-244;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSSSGLLGYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSSSGLLGYYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDVETPATSADNHVTMMVDQEVINKASNKIRLEKGLYQIKIY 120

Db 61 QSAIWSGFIKVKSDSEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGLVQIKQY 120  
QY 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLQPELKOKSNSRKRKSTSGPTVPDRDN 180  
Db 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLQPELKOKSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240  
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWATWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWATWGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNCNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNCNL 540  
QY 541 QYQKDIITEFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKNAKNMILIRDRFFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKNAKNMILIRDRFFHYDR 600  
QY 601 NNIAVGADES VVKAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADES VVKAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

RESULT 6  
AAM51488  
ID AAM51488 standard; protein; 735 AA.  
XX  
AC AAM51488;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Anthrax PA mutant D425A.  
XX  
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.  
XX  
OS Bacillus anthracis.  
OS  
XX  
PH Key Location/Qualifiers  
FT Misc-difference 425 /note= "Wild-type Asp substituted by Ala"  
XX  
XX WO200182788-A2.  
XX  
XX 08-NOV-2001.  
XX  
XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX Collier RJ, Sellman BR;  
XX WPI; 2002-017725/02.  
XX Protecting humans against anthrax using mutant B groups (anthrax  
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX  
XX Claim 4; Page; 77pp; English.  
XX  
XX The invention relates to antibacterial agents comprising mutant forms of  
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX moiety is anthrax protective antigen (PA) and using these mutants or  
XX compositions of them for protecting against Bacillus anthracis infections  
XX in humans, especially as vaccines. Note: The present sequence is not  
XX given in the specification but is derived from the Bacillus anthracis  
XX wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
XX Sequence 735 AA;  
XX  
XX Query Match 99.8%; Score 3767; DB 5; Length 735;  
XX Best Local Similarity 99.9%; Pred. No. 1.7e-244;  
XX Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENLLNESSSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENLLNESSSSQGLLGYYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDSEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGLVQIKQY 120  
Db 61 QSAIWSGFIKVKSDSEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGLVQIKQY 120  
QY 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLQPELKOKSNSRKRKSTSGPTVPDRDN 180  
Db 121 QRENTEKGLDFKLWYDSQNKKEVSSDNLQPELKOKSNSRKRKSTSGPTVPDRDN 180  
QY 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240  
Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWATWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWATWGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNIAATYFNGRVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNCNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNCNL 540  
QY 541 QYQKDIITEFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKNAKNMILIRDRFFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKNAKNMILIRDRFFHYDR 600  
QY 601 NNIAVGADES VVKAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADES VVKAHREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660

661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSNG 720  
 QY  
 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSNG 720  
 Db  
 721 IKKILIFSKKGYEIG 735  
 QY  
 721 IKKILIFSKKGYEIG 735  
 Db  
 RESULT 7  
 AAM51493  
 ID AAM51493 standard; protein; 735 AA.  
 XX  
 AC AAM51493;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Anthrax PA mutant K397D/D425K.  
 XX  
 KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
 XX  
 OS Bacillus anthracis.  
 XX  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"  
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"  
 FT  
 XX  
 PN WC200182788-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 XX 04-MAY-2001; 2001WO-US014372.  
 XX  
 XX 04-MAY-2000; 2000US-0201800P.  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 XX  
 PI Collier RJ, Sellman BR;  
 XX  
 XX WPI; 2002-017725/02.  
 XX  
 PT Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.  
 XX  
 PS Claim 4; Page; 77pp; English.  
 XX  
 XX The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
 XX  
 SQ Sequence 735 AA;  
 Query March 99.8%; Score 3767; DB 5; Length 735;  
 Best Local Similarity 99.9%; Pred. No. 1.7e-244;  
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
 Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
 QY 61 QSAIWGFIKVKSGDEYTFATSNADNHTVMWDDQEVINKASNNKIRLEKGRLYQIKIY 120

Db 61 QSAIWGFIKVKSGDEYTFATSNADNHTVMWDDQEVINKASNNKIRLEKGRLYQIKIY 120  
 QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPKLKQKSSNRKCRSTASGPTVPDRN 180  
 Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPKLKQKSSNRKCRSTASGPTVPDRN 180  
 QY 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNHEKKGITKYKSPKXWSTASDPYSDFEKT 240  
 Db 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNHEKKGITKYKSPKXWSTASDPYSDFEKT 240  
 QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTGRHT 300  
 Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTGRHT 300  
 QY 301 SEVHGNAEVEHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 Db 301 SEVHGNAEVEHASFDDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420  
 Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPPSKNLAPIA 420  
 QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATATYFNGRVRVDTGSKWSEV 480  
 Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATATYFNGRVRVDTGSKWSEV 480  
 QY 481 LPOIETTARIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLKEALKIATFGNEPENG 540  
 Db 481 LPOIETTARIIFNGKDLNVERRIAAVNPSPDELTTPKDMTLKEALKIATFGNEPENG 540  
 QY 541 QYQKDTTEFDNFDDQOTSONIKNQLAELNATYIVLDKIKLNKMTILIRDKRPHYDR 600  
 Db 541 QYQKDTTEFDNFDDQOTSONIKNQLAELNATYIVLDKIKLNKMTILIRDKRPHYDR 600  
 QY 601 NNIAVGADSVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
 Db 601 NNIAVGADSVVKEAHEVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
 QY 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSNG 720  
 Db 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPNYKVVAVTKENTIIINPSENGDTSNG 720  
 QY 721 IKKILIFSKKGYEIG 735  
 Db 721 IKKILIFSKKGYEIG 735  
 RESULT 8  
 AAY56959  
 ID AAY56959 standard; protein; 736 AA.  
 XX  
 AC AAY56959;  
 XX  
 DT 25-APR-2000 (first entry)  
 XX  
 DE B. anthracis MAT-PA protein.  
 XX  
 KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;  
 KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.  
 XX  
 OS Bacillus anthracis.  
 XX  
 PN WO200002522-A2.  
 XX  
 PD 20-JAN-2000.  
 XX  
 PF 09-JUL-1999; 99WO-US015568.  
 XX  
 PR 10-JUL-1998; 98US-0092416P.  
 XX  
 PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.  
 XX  
 PI Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;

XX WPI; 2000-182165/16.  
DR N-PSDB; AAZ56875.  
XX Recombinant DNA construct useful as vaccines for anthrax, in producing  
PT host cells for analyzing the drugs and agents inhibiting anthrax.  
XX  
PS Disclosure; Page 34; 35pp; English.  
XX  
CC The invention provides a recombinant DNA construct that comprises a  
CC vector and at least one nucleic acid (or its fragment) encoding a  
CC combination of Bacillus anthracis proteins, selected from protective  
CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA  
CC with its secretory signals replaced with those of tissue plasminogen  
CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine  
CC for anthrax and in producing infectious alpha virus particles. These  
CC particles, expressing the B. anthracis proteins are useful also as  
CC vaccines for anthrax. Host cells transformed with the construct are  
CC useful for analyzing the effectiveness of drugs and agents that inhibit  
CC anthrax or B. anthracis proteins. The present sequence represents a B.  
CC anthracis MAT-PA protein  
XX  
SQ Sequence 736 AA;  
  
Query Match 99.8%; Score 3767; DB 3; Length 736;  
Best Local Similarity 99.9%; Pred. No. 1.7e-244;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
Db 2 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 61  
  
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHTVMWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 62 QSAIWSGFIKVKKSDDEYTFATSDADNHTVMWDDQEVINKASNSKIRLEKGRLYQIKIY 121  
  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRSTSGPTVPDRDN 180  
Db 122 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRSTSGPTVPDRDN 181  
  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPFKWSTASDPYSDFEKT 240  
Db 182 DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPFKWSTASDPYSDFEKT 241  
  
QY 241 GRIDKNVSEARPLVAAPVIVHVDENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 242 GRIDKNVSEARPLVAAPVIVHVDENILSKNEDQSTQNTDSETRISKNTSRTHT 301  
  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 302 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 361  
  
QY 361 NANIRYVNTGTAPIYVLPSTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 362 NANIRYVNTGTAPIYVLPSTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 421  
  
QY 421 LNAQKQSPPTIMYNOFLEKTKQLRLDTPDVGNATYNFENGRVVRTGNSWSEV 480  
Db 422 LNAQKQSPPTIMYNOFLEKTKQLRLDTPDVGNATYNFENGRVVRTGNSWSEV 481  
  
QY 481 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 540  
Db 482 LPQIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPNGNL 541  
  
QY 541 QYQKQITEFDNFDDQTSQNIKNQALAEANATNIVYVLDKIKLNKQNLILRDKRPHYDR 600  
Db 542 QYQKQITEFDNFDDQTSQNIKNQALAEANATNIVYVLDKIKLNKQNLILRDKRPHYDR 601  
  
QY 601 NNIAGADESVKAEHREVINSTEGLLLNIDKIDKILSGYIVETEDTGLKEVINDRY 660  
Db 602 NNIAGADESVKAEHREVINSTEGLLLNIDKIDKILSGYIVETEDTGLKEVINDRY 661  
  
QY 661 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVVAVTKENTIIINPSENGDTSTNG 720

Db 662 DMLNISSLRQDGKTFIDFKYNDKLPYISNPYKVVAVTKENTIIINPSENGDTSTNG 721  
QY 721 IKKILIFSCKGVEIG 735  
Db 722 IKKILIFSCKGVEIG 736  
  
RESULT 9  
AAZ56960  
ID AAZ56960 standard; protein; 763 AA.  
XX  
XX AAZ56960;  
XX  
DT 25-APR-2000 (first entry)  
XX  
DE B. anthracis TPA-PA protein.  
XX  
KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;  
KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.  
XX  
OS Bacillus anthracis.  
XX  
XX WO200002522-A2.  
XX  
PD 20-JAN-2000.  
XX  
PF 09-JUL-1999; 99WO-US015568.  
XX  
PR 10-JUL-1998; 98US-0092416P.  
XX  
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.  
XX  
XX Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;  
PI WPI; 2000-182165/16.  
XX  
DR N-PSDB; AAZ56876.  
XX  
PT Recombinant DNA construct useful as vaccines for anthrax, in producing  
FT host cells for analyzing the drugs and agents inhibiting anthrax.  
XX  
PS Disclosure; Page 32; 35pp; English.  
XX  
CC The invention provides a recombinant DNA construct that comprises a  
CC vector and at least one nucleic acid (or its fragment) encoding a  
CC combination of Bacillus anthracis proteins, selected from protective  
CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA  
CC with its secretory signals replaced with those of tissue plasminogen  
CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine  
CC for anthrax and in producing infectious alpha virus particles. These  
CC particles, expressing the B. anthracis proteins are useful also as  
CC vaccines for anthrax. Host cells transformed with the construct are  
CC useful for analyzing the effectiveness of drugs and agents that inhibit  
CC anthrax or B. anthracis proteins. The present sequence represents a B.  
CC anthracis TPA-PA protein  
XX  
SQ Sequence 763 AA;  
  
Query Match 99.8%; Score 3767; DB 3; Length 763;  
Best Local Similarity 99.9%; Pred. No. 1.8e-244;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
Db 29 EVKQENRLNSESQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 88  
  
QY 61 QSAIWSGFIKVKKSDDEYTFATSDADNHTVMWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 89 QSAIWSGFIKVKKSDDEYTFATSDADNHTVMWDDQEVINKASNSKIRLEKGRLYQIKIY 148  
  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRSTSGPTVPDRDN 180  
Db 149 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKQSSNSRKRSTSGPTVPDRDN 208



QY 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYSDFEKVT 240  
 Db 209 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYSDFEKVT 268  
 QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 300  
 Db 269 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 328  
 QY 301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 Db 329 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 388  
 QY 361 NANIRYVNTGTAPIYVNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
 Db 389 NANIRYVNTGTAPIYVNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 448  
 QY 421 LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 480  
 Db 449 LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 508  
 QY 481 LPQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540  
 Db 509 LPQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 568  
 QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLIRDKRFHYDR 600  
 Db 569 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLIRDKRFHYDR 628  
 QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
 Db 629 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 688  
 QY 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNYAVTKENTIINPSENGDTSTNG 720  
 Db 689 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNYAVTKENTIINPSENGDTSTNG 748  
 QY 721 IKKILIFSKKGYEIG 735  
 Db 749 IKKILIFSKKGYEIG 763  
 RESULT 10  
 AAY56958  
 ID AAY56958 standard; protein; 764 AA.  
 XX  
 AC AAY56958;  
 DT 25-APR-2000 (first entry)  
 XX  
 DE B. anthracis protective antigen (PA) protein.  
 XX  
 KW Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;  
 KW tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.  
 XX  
 OS Bacillus anthracis.  
 XX  
 FN WO200002522-A2.  
 XX  
 PD 20-JAN-2000.  
 XX  
 PF 09-JUL-1999; 99WO-US015568.  
 XX  
 PR 10-JUL-1998; 98US-0092416P.  
 XX  
 PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.  
 XX  
 PI Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;  
 XX  
 DR WPI; 2000-182165/16.  
 DR N-PSDB; AAZ56874.  
 XX  
 PT Recombinant DNA construct useful as vaccines for anthrax, in producing

PT host cells for analyzing the drugs and agents inhibiting anthrax.  
 XX Disclosure; Page 33; 35pp; English.  
 XX  
 CC The invention provides a recombinant DNA construct that comprises a  
 CC vector and at least one nucleic acid (or its fragment) encoding a  
 CC combination of Bacillus anthracis proteins, selected from protective  
 CC antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA  
 CC with its secretory signals replaced with those of tissue plasminogen  
 CC activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine  
 CC for anthrax and in producing infectious alpha virus particles. These  
 CC particles, expressing the B. anthracis proteins are useful also as  
 CC vaccines for anthrax. Host cells transformed with the construct are  
 CC useful for analyzing the effectiveness of drugs and agents that inhibit  
 CC anthrax or B. anthracis proteins. The present sequence represents a B.  
 CC anthracis PA protein  
 XX  
 SQ Sequence 764 AA;  
 Query Match 99.8%; Score 3767; DB 3; Length 764;  
 Best Local Similarity 99.9%; Pred. No. 1.8e-244;  
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EVKQENRLLNESSESSQGLLGYYFSDLNPOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
 Db 30 EVKQENRLLNESSESSQGLLGYYFSDLNPOAPMVVTSSTGDLSPSSSELENIPSENQYF 89  
 QY 61 QSAIWSGFIVKKSDEYTPATSADNHVTVMVDDQEVINKASNSNKIRLEKRLYQIKQY 120  
 Db 90 QSAIWSGFIVKKSDEYTPATSADNHVTVMVDDQEVINKASNSNKIRLEKRLYQIKQY 149  
 QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNRKRSAGTSAGTVDPDRN 180  
 Db 150 QRENPTKGLDFKLYWTDSONKKEVSSDNLQJPELKQKSSNRKRSAGTSAGTVDPDRN 209  
 QY 181 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYSDFEKVT 240  
 Db 210 DGIPLSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYSDFEKVT 269  
 QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 300  
 Db 270 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQMTDSETRISKVTSRTH 329  
 QY 301 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 Db 330 SEVHGNAEVAHAFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 389  
 QY 361 NANIRYVNTGTAPIYVNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
 Db 390 NANIRYVNTGTAPIYVNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 449  
 QY 421 LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 480  
 Db 450 LNAQKDFSTPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSWSEV 509  
 QY 481 LPQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540  
 Db 510 LPQIOETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 569  
 QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLIRDKRFHYDR 600  
 Db 570 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMILLIRDKRFHYDR 629  
 QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
 Db 630 NNIAVGADESVMKEAHREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 689  
 QY 661 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNYAVTKENTIINPSENGDTSTNG 720  
 Db 690 DMLNSSLRODQKTFIDFKYNDKPLIYISNPNKYVNYAVTKENTIINPSENGDTSTNG 749  
 QY 721 IKKILIFSKKGYEIG 735  
 |||||

Db	750	IKKILIFSKKGYEIG 764	
Db	30	EVKQENRLNSESSESSQGLLYYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 89	
QY	61	QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQBYINKASNSNKIRLEKGRLYQIKIY 120	
Db	90	QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQBYINKASNSNKIRLEKGRLYQIKIY 149	
QY	121	QRENPTKGLDFKLWYTDSONKKEVISSDNQLPELQKXSSNSRKRKSTAGTVPDRDN 180	
Db	150	QRENPTKGLDFKLWYTDSONKKEVISSDNQLPELQKXSSNSRKRKSTAGTVPDRDN 209	
QY	181	DGIPDSLEVEGYTVDVKNKTEFLSPWISNHEKKGTLTKYKSSPEKWTASDPYDPEKVT 240	
Db	210	DGIPDSLEVEGYTVDVKNKTEFLSPWISNHEKKGTLTKYKSSPEKWTASDPYDPEKVT 269	
QY	241	GRIDKNVSPARHPLVAAYPIVHVDMENILSKNEDQSTONTDSETRTISKNTSTRTHT 300	
Db	270	GRIDKNVSPARHPLVAAYPIVHVDMENILSKNEDQSTONTDSETRTISKNTSTRTHT 329	
QY	301	SEVHGNAEVAHSPFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360	
Db	330	SEVHGNAEVAHSPFDIGGSVSGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 389	
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPKSLAPIA 420	
Db	390	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYYPKSLAPIA 449	
QY	421	LNAQDFSTPTITMNYNQFLEKTKQLRLDTQVYGNIAFYENFENGVRVDTGSKNWEV 480	
Db	450	LNAQDFSTPTITMNYNQFLEKTKQLRLDTQVYGNIAFYENFENGVRVDTGSKNWEV 509	
QY	481	LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540	
Db	510	LPOIQTETARIIFNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 569	
QY	541	QYQKDIETFDNPDQOSTQNIKNQLAELNATNIYVLDKIKLNAKVNILIRKRFHYDR 600	
Db	570	QYQKDIETFDNPDQOSTQNIKNQLAELNATNIYVLDKIKLNAKVNILIRKRFHYDR 629	
QY	601	NNIAVGADESUVKEAHRVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660	
Db	630	NNIAVGADESUVKEAHRVINSSTEGILLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 689	
QY	661	DMLNISSLRQDKTFIDFKYNDKLPYIYISNPNYKVNVAVTENTIIINSENGDTSTNG 720	
Db	690	DMLNISSLRQDKTFIDFKYNDKLPYIYISNPNYKVNVAVTENTIIINSENGDTSTNG 749	
QY	721	IKKILIFSKKGYEIG 735	
Db	750	IKKILIFSKKGYEIG 764	
Db	RESULT 12		
Db	ADE65872		
ID	ADE65872	standard; protein; 857 AA.	
XX	AC	ADE65872;	
XX	DT	29-JAN-2004 (first entry)	
XX	DE	Bacillus anthracis PA antigen.	
XX	KW	Antibacterial; Vaccine; immune response; Bacillus anthracis.	
XX	OS	Bacillus anthracis.	
XX	PN	WO2003087378-A1.	
XX	PD	23-OCT-2003.	
XX	PF	11-APR-2003; 2003WO-GB001553.	
XX	PR	11-APR-2002; 2002US-0371416P.	
XX	XX		

Query Match 99.8%; Score 3767; DB 4; Length 764;  
Best Local Similarity 99.9%; Pred. No. 1.8e-244;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
|||||

Db 742 DMLNTSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKNTIINPSENGDTSTNG 801

Qy 721 IKKILFSSKGYBIG 735

Db 802 IKKILFSSKGYBIG 816

RESULT 13

AAE18289

AC AAE18289;

DT 07-MAY-2002 (first entry)

XX Bacillus subtilis protective antigen (PA) domain 1+2+3+4.

XX Immunogenic reagent; immune response; protective antigen; PA; vaccine;

XX Bacillus anthracis infection; antibacterial.

XX Bacillus subtilis.

XX Key Location/Qualifiers

FT Misc-difference 285 /note= "Encoded by GAA"

XX WO200204646-A1.

XX 17-JAN-2002.

XX 06-JUL-2001; 2001WO-GB003065.

XX 08-JUL-2000; 2000GB-00016702.

XX (MINA ) UK SEC FOR DEFENCE.

XX Williamson ED, Miller J, Walker NJ, Baillie LMJ, Holden PT;

XX Flick-Smith HC, Bullifent HL, Titball RW, Topping AW;

XX WPI; 2002-171720/22.

XX N-PSDB; AAD29121.

XX New immunogenic reagent having a polypeptide of the full length

XX Protective Antigen of Bacillus anthracis, useful for treating B.

XX anthracis infection or in preparing a medicament for the prophylaxis or

XX treatment of the infection.

XX Example 5; Fig 3; 40pp; English.

XX The present invention relates to an immunogenic reagent, which produces

XX an immune response that is protective against Bacillus anthracis. The

XX reagent comprises one or more polypeptides which together represent up to

XX three domains of the full length Protective Antigen (PA) of Bacillus

XX anthracis or variants of these, and at least one of the domains comprises

XX domain 1 or domain 4 of PA or its variant. The invention is used as a

XX vaccine. The immunogenic reagent is useful in the preparation of a

XX medicament for the prophylaxis or treatment of B. anthracis infection.

XX The present sequence is Bacillus subtilis protective antigen domain

XX 1+2+3+4

SQ Sequence 735 AA;

Query Match 99.8%; Score 3764; DB 5; Length 735;

Best Local Similarity 99.7%; Pred. No. 2.8e-244;

Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Qy 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKQY 120

Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKQY 120

PA (POWD-) POWDERJECT RES LTD.

XX Schmaljohn C, Fuller J;

XX WPI; 2003-877105/81.

XX N-PSDB; ADE56871.

XX New polynucleotide vaccine composition comprising a nucleic acid sequence

XX that encodes a Bacillus anthracis antigen, useful for eliciting a

XX protective immune response against Bacillus anthracis.

XX Example 1; SEQ ID NO 4; 65pp; English.

XX The present invention relates to a new polynucleotide vaccine composition

XX comprising a nucleic acid sequence that encodes a Bacillus anthracis

XX antigen and that is operatively linked to a promoter suitable for

XX expression of the antigen in a mammalian cell. The polynucleotide vaccine

XX composition is useful for eliciting a protective immune response against

XX Bacillus anthracis. The present sequence represents Bacillus anthracis PA

XX antigen.

XX Sequence 857 AA;

Query Match 99.8%; Score 3767; DB 7; Length 857;

Best Local Similarity 99.9%; Pred. No. 2.2e-244;

Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60

Db 82 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 141

Qy 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKQY 120

Db 142 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKQY 201

Qy 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQIPELKQSSNRKRGSTAGTVPDRDN 180

Db 202 QRENPTKGLDFKLYWTDSONKEVSSDNLQIPELKQSSNRKRGSTAGTVPDRDN 261

Qy 181 DGIPDSLEVEGYTVDVNKKFTLSPWISNIHEKKGLTKYKSSPEKSTASDPYDPEKVT 240

Db 262 DGIPDSLEVEGYTVDVNKKFTLSPWISNIHEKKGLTKYKSSPEKSTASDPYDPEKVT 321

Qy 241 GRIDKNVSPPEARHPVLAAYPIVHVDMENIILSKNEDQSTONTSETRTSKNTSTSRTH 300

Db 322 GRIDKNVSPPEARHPVLAAYPIVHVDMENIILSKNEDQSTONTSETRTSKNTSTSRTH 381

Qy 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Db 382 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 441

Qy 361 NANIRYVNTGTAPVYVLPVTSVLYGKQNTLAKAKENQSLQILAPNNYPSKNLAPIA 420

Db 442 NANIRYVNTGTAPVYVLPVTSVLYGKQNTLAKAKENQSLQILAPNNYPSKNLAPIA 501

Qy 421 LNAQKDFSSSTITWNYQFLEKTKQIRLDTDOVYGNATYFENGVRVVDTCGNWSEV 480

Db 502 LNAQKDFSSSTITWNYQFLEKTKQIRLDTDOVYGNATYFENGVRVVDTCGNWSEV 561

Qy 481 LPQIQETTARIIFNGKDLNLYVERIAAVNPSDPLETTKPDWTLKEALKIAGFNEPENG 540

Db 562 LPQIQETTARIIFNGKDLNLYVERIAAVNPSDPLETTKPDWTLKEALKIAGFNEPENG 621

Qy 541 QYQGHDIETFPNFDQOTSQNIKNQALNATNYVLDKTKNAKNILIRDKRPHYDR 600

Db 622 QYQGHDIETFPNFDQOTSQNIKNQALNATNYVLDKTKNAKNILIRDKRPHYDR 681

Qy 601 NNIAVGADSVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIETDEGLKEVINDRY 660

Db 682 NNIAVGADSVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIETDEGLKEVINDRY 741

Qy 661 DMLNTSSLRQDGKTFIDFKKYNKLPYISNPNYKVNVAVTKNTIINPSENGDTSTNG 720

Sutton JM, Shone CC;  
WPI; 2003-167247/16.  
Conjugate for modulating cell survival and cell growth, modulating release of inflammatory mediator from cells, comprises injected bacterial effector protein and a carrier that targets the protein to target cell.  
Example 12; Page 122-125; 130pp; English.  
The invention relates to a conjugate comprising an injected bacterial effector protein and a carrier that targets the effector protein to a target cell. Pharmaceutical composition of the invention is useful for a treatment selected from promoting or inhibiting survival of cells; preventing and reversing damage to cells; killing cells; promoting or inhibiting the growth of cells; apoptosis, release of an inflammatory mediator from cells, division of cells and treating intracellular infection and regulating nitric oxide release from cells. The invention is useful in the manufacture of a medicament for treating a neuronal cell, for intracellular infection, for interfering with intracellular trafficking, for modulating expression of cell-surface markers and for inhibiting secretion from cells. The invention is also useful for treating Prion disease, Alzheimer' disease and wide range of disorders including muscle spasms such as blepharospasm, torticollis and hypersecretion disorders such as chronic obstructive pulmonary disease (COPD), bronchitis and asthma. The present sequence is Bacillus anthracis protective antigen. This sequence is used in the exemplification of the invention  
Sequence 764 AA;  
Query Match 99.8%; Score 3764; DB 6; Length 764;  
Best Local Similarity 99.7%; Pred. No. 2.9e-244;  
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENLLNESESSQGLLYGYPFDLNFQAPMVTTSITGDLSPSELENIPSENYQYF 60  
Db 30 EVKQENLLNESESSQGLLYGYPFDLNFQAPMVTTSITGDLSPSELENIPSENYQYF 89  
QY 61 QSAIWSGFIVKKSDEYTFATSDNHYTMVDDQEVINKASNNKIRLEKGRLYQIKIY 120  
Db 90 QSAIWSGFIVKKSDEYTFATSDNHYTMVDDQEVINKASNNKIRLEKGRLYQIKIY 149  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSISSDNLQPELKQKSSNRKKESTAGPTVPDRDN 180  
Db 150 QRENPTKGLDFKLYWTDSONKKEVSISSDNLQPELKQKSSNRKKESTAGPTVPDRDN 209  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 240  
Db 210 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 269  
QY 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNSTSRHT 300  
Db 270 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNSTSRHT 329  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 360  
Db 330 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 389  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 390 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 449  
QY 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGSNWSEV 480  
Db 450 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGSNWSEV 509  
QY 481 LPQIOETTARIIFNGKOLNLYVERIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPKNL 540  
Db 510 LPQIOETTARIIFNGKOLNLYVERIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPKNL 569  
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNYTVLDKIKLNAKONILLRDKRPHYDR 600

QY 121 QRENPTKGLDFKLYWTDSONKKEVSISSDNLQPELKQKSSNRKKESTAGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSISSDNLQPELKQKSSNRKKESTAGPTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKNSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSEITISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAEITGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITWYNOFLEKTKQLRLDQVYGNATYFNENGRVVRVDTGSNWSEV 480  
QY 481 LPQIOETTARIIFNGKOLNLYVERIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPKNL 540  
Db 481 LPQIOETTARIIFNGKOLNLYVERIAAVNPSDPLETTKPDMTLKEALKIAFGFNEPKNL 540  
QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNYTVLDKIKLNAKONILLRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNYTVLDKIKLNAKONILLRDKRPHYDR 600  
QY 601 NNTAVGADSEVVKAEHREVNSTEGLLNIDKIRKILSGYIWEDETEGLKEVINDRY 660  
Db 601 NNTAVGADSEVVKAEHREVNSTEGLLNIDKIRKILSGYIWEDETEGLKEVINDRY 660  
QY 661 DMLNSSLRODGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODGTFFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIPSKKGYEIG 735  
Db 721 IKKILIPSKKGYEIG 735  
RESULT 14  
AAE35717  
ID AAE35717 standard; protein; 764 AA.  
XX AAE35717;  
XX 17-JUN-2003 (first entry)  
XX Bacillus anthracis protective antigen.  
XX Apoptosis; therapy; inflammatory mediator; intracellular trafficking;  
XX infection; Prion disease; Alzheimer' disease; hypersecretion disorder;  
XX muscle spasm; COPD; bronchitis; chronic obstructive pulmonary disease;  
XX torticollis; blepharospasm; asthma; protective antigen.  
XX Bacillus anthracis.  
XX W0200296467-A2.  
XX 05-DEC-2002.  
XX 21-MAY-2002; 2002WO-GB002384.  
XX 24-MAY-2001; 2001GB-00012687.  
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.  
XX

Db 570 QYQKDIITFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKMNILIRDKRPHYDR 629  
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
Db 630 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 689  
QY 661 DMLNISLRODQGTFFDFKKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 690 DMLNISLRODQGTFFDFKKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 749  
QY 721 IKKILIFSCKGYEIG 735  
Db 750 IKKILIFSCKGYEIG 764

RESULT 15  
AAM51487  
ID AAM51487 standard; protein; 735 AA.  
XX AC AAM51487;  
XX DT 01-FEB-2002 (first entry)  
XX DE Anthrax PA mutant K397Q.  
XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX OS Bacillus anthracis.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Misc-difference 397 /note= "Wild-type Lys substituted by Gln"  
XX PN W0200182788-A2.  
XX PD 08-NOV-2001.  
XX PF 04-MAY-2001; 2001WO-US014372.  
XX PR 04-MAY-2000; 2000US-0201800P.  
XX PA (HARD ) HARVARD COLLEGE.  
XX PI Collier RJ, Sellman BR;  
XX DR WPI; 2002-017725/02.  
XX PT Protecting humans against anthrax using mutant B groups (anthrax  
XX PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
XX PS anthracis.  
XX PS Claim 4; Page; 77pp; English.  
XX CC The invention relates to antibacterial agents comprising mutant forms of  
XX CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
XX CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
XX CC moiety is anthrax protective antigen (PA) and using these mutants or  
XX CC compositions of them for protecting against Bacillus anthracis infections  
XX CC in humans, especially as vaccines. Note: The present sequence is not  
XX CC given in the specification but is derived from the Bacillus anthracis  
XX CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX SQ Sequence 735 AA;

Query Match 99.7%; Score 3763; DB 5; Length 735;  
Best Local Similarity 99.7%; Pred. No. 3.2e-244;  
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSESSGCLIGYFSDLNFCQAPMVVTSSTTGDLISPSSELENIPSENQYF 60  
Db 1 EVKQENLLNESSESSGCLIGYFSDLNFCQAPMVVTSSTTGDLISPSSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLSEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSNKIRLSEKGRLYQIKIY 120  
QY 121 QRENPEKGLDFKLYMTDSQNKKEVLSNQLPELKOKSSNSRKRKSTASAGTVPDRDN 180  
Db 121 QRENPEKGLDFKLYMTDSQNKKEVLSNQLPELKOKSSNSRKRKSTASAGTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSDPEKVT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDHNIILSKNEDQSTQNTDSETRTISKNTSTRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDHNIILSKNEDQSTQNTDSETRTISKNTSTRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYNTGTAPIYVLPITTSVLGKNQTLATIKAKENOLSOILAPNNYYPKNIAPIA 420  
Db 361 NANIRYNTGTAPIYVLPITTSVLGKNQTLATIKAKENOLSOILAPNNYYPKNIAPIA 420  
QY 421 LNAQKOPSSPTITWYVNFQFLELEKTKQLRLDTPQVYGNIAIYNFENGVRVDTGSNWEV 480  
Db 421 LNAQKOPSSPTITWYVNFQFLELEKTKQLRLDTPQVYGNIAIYNFENGVRVDTGSNWEV 480  
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENGIL 540  
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAFGNEPENGIL 540  
QY 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKMNILIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNAKMNILIRDKRPHYDR 600  
QY 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHREVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
QY 661 DMLNISLRODQGTFFDFKKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISLRODQGTFFDFKKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735

Search completed: May 3, 2004, 19:35:55  
Job time : 52.1208 secs



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-4

Query Match 99.8%; Score 3767; DB 1; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.3e-256;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTSQNKKEVTSNDLQPELKOKSSNSRKRSTAGTPVDPDRN 180  
DB 121 QRENTEKGLDFKLYWTSQNKKEVTSNDLQPELKOKSSNSRKRSTAGTPVDPDRN 180  
QY 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240  
DB 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
DB 301 SEVHGNAEVHASFFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITWYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
DB 421 LNAQKDFSSPTITWYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILIRDKRFHYDR 600  
DB 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILIRDKRFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDKIRKILSGYIIEDETEGLKEVINDRY 660  
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDKIRKILSGYIIEDETEGLKEVINDRY 660

DB 1 EVKQENLLNESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120  
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTSQNKKEVTSNDLQPELKOKSSNSRKRSTAGTPVDPDRN 180  
DB 121 QRENTEKGLDFKLYWTSQNKKEVTSNDLQPELKOKSSNSRKRSTAGTPVDPDRN 180  
QY 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240  
DB 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
DB 241 GRIDKNVSPPEARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
DB 301 SEVHGNAEVHASFFDGGSVSAGFSNSSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITWYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
DB 421 LNAQKDFSSPTITWYNNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540  
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILIRDKRFHYDR 600  
DB 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNYTVLDKIKLAKNNILIRDKRFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDKIRKILSGYIIEDETEGLKEVINDRY 660  
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDKIRKILSGYIIEDETEGLKEVINDRY 660  
QY 661 DMLNLSLRQGGTKTFIDFKYNDKPLYSNPKNYKVVAVTKENTIIINPSENGDTSTNG 720  
DB 661 DMLNLSLRQGGTKTFIDFKYNDKPLYSNPKNYKVVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
DB 721 IKKILIFSCKGYEIG 735

RESULT 2  
US-082-849B-4  
Sequence 4, Application US/08082849B  
Parent No. 5677274  
GENERAL INFORMATION:  
APPLICANT: Lepplia, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
TITLE OF INVENTION: Related Methods  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

661 DMLNLSLRQDGTFFIDFKKNDKLPYISNPNKVNVAVTKENTIIINSENGDTSTNG 720  
661 DMLNLSLRQDGTFFIDFKKNDKLPYISNPNKVNVAVTKENTIIINSENGDTSTNG 720  
721 IKKILIFSKGYEIG 735  
721 IKKILIFSKGYEIG 735  
RESULT 3  
PCT-US94-01624-4  
Sequence 4, Application PC/TUS9401624  
GENERAL INFORMATION:  
APPLICANT: Leppia, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
RELATED METHODS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW  
STREET: Steuart Street Tower, 20th Floor, One Market  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/01624  
FILING DATE: June 25, 1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01624-4  
Query Match 99.8%; Score 3767; DB 5; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.3e-256;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSTTGLSTPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSQGLGYFSDLNFAQPMVVTSTTGLSTPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDYTFATSDADNHVTWVDDQEVINKASNKRIRLEKGRLYQIKQY 120  
Db 61 QSAIWSGFIKVKSDYTFATSDADNHVTWVDDQEVINKASNKRIRLEKGRLYQIKQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVTSNQLQPELKQKSSNSRKRSTAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKEVTSNQLQPELKQKSSNSRKRSTAGTVPDRDN 180  
QY 181 DGIPTDSLEVEGYTVDVNKKETFLSPWISNHEKKGITKYKSSPEKWTASDPYDFEYKT 240  
Db 181 DGIPTDSLEVEGYTVDVNKKETFLSPWISNHEKKGITKYKSSPEKWTASDPYDFEYKT 240

241 GRIDKNVSPARHPVAAPIVHVDMENIILSKNEDOSTQNTDSETRISKNSTSRHT 300  
241 GRIDKNVSPARHPVAAPIVHVDMENIILSKNEDOSTQNTDSETRISKNSTSRHT 300  
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWASTMGLNTADTARL 360  
301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERWASTMGLNTADTARL 360  
361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDQVYGNIAIYNFENGRVVDVTGSNWSEV 480  
421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDQVYGNIAIYNFENGRVVDVTGSNWSEV 480  
481 LPQIQETIARIIFNGKDLNVERIAAANPSPLETTKPDMTLKEALKIARFENPENG 540  
481 LPQIQETIARIIFNGKDLNVERIAAANPSPLETTKPDMTLKEALKIARFENPENG 540  
541 QYQKDKITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKMNILIRDKRFHYDR 600  
541 QYQKDKITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKMNILIRDKRFHYDR 600  
601 NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
601 NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
661 DMLNLSLRQDGTFFIDFKKNDKLPYISNPNKVNVAVTKENTIIINSENGDTSTNG 720  
661 DMLNLSLRQDGTFFIDFKKNDKLPYISNPNKVNVAVTKENTIIINSENGDTSTNG 720  
721 IKKILIFSKGYEIG 735  
721 IKKILIFSKGYEIG 735  
RESULT 4  
US-08-021-601-12  
Sequence 12, Application US/08021601  
Patent No. 5591631  
GENERAL INFORMATION:  
APPLICANT: Leppia, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Nichols, Peter J.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
RELATED METHODS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Needle & Rosenberg, P.C.  
STREET: 133 Carnegie Way, Suite 400  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/021,601  
FILING DATE: 19930212  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Spratt, Gwendolyn D.  
REGISTRATION NUMBER: 36,016  
REFERENCE/DOCKET NUMBER: 1414.057  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404/688-0770  
TELEFAX: 404/688-9880



```
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-021-601-12

Query Match      98.7%; Score 3722.5; DB 1; Length 903;
Best Local Similarity 99.0%; Pred. No. 2.3e-253;
Matches 727; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 EVKQENRLNESSSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSPKMWSTASDPYSDFEKT 240
DB 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSPKMWSTASDPYSDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTHT 300
DB 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTIMWYNOFLEKTKQLRLDQVYGNATYFNFENGVRVDTGSGWSEV 480
DB 421 LNAQKDFSSPTIMWYNOFLEKTKQLRLDQVYGNATYFNFENGVRVDTGSGWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETKPDMTLKEALKIAFGNFPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETKPDMTLKEALKIAFGNFPNGNL 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRFHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKKNILIRDKRFHYDR 600
QY 601 NNIAVGADESVMKAEHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKAEHREVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNKYKNVYAVTKNTIINPSENGDTSTNG 720
DB 661 DMLNSSLRODGTFFIDFKYNDKPLIYISNPNKYKNVYAVTKNTIINPSENGDTSTNG 720
QY 721 IKKILKVVVLGKKG 731
DB 721 IKKILKVVVLGKKG 734
```

```
RESULT 5
US-08-082-849B-12
; Sequence 12, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Klimpel, Kurt R.
; APPLICANT: Arora, Naveen
```

```
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,849B
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,601
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-161-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 903 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-082-849B-12
```

```
Query Match      98.7%; Score 3722.5; DB 1; Length 903;
Best Local Similarity 99.0%; Pred. No. 2.3e-253;
Matches 727; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 EVKQENRLNESSSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
DB 1 EVKQENRLNESSSQGLLYGFFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180
QY 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSPKMWSTASDPYSDFEKT 240
DB 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGITKYKSPKMWSTASDPYSDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTHT 300
DB 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSSTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYNYLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTIMWYNOFLEKTKQLRLDQVYGNATYFNFENGVRVDTGSGWSEV 480
```

Db 421 LNAQDDFSPTITMNNYQFLELEKTKQLRLDQVYGNIAVNFENGVRVDTGNNWSEV 480  
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAANPSPLETTKPDMLKEALKIAGFNEPENGNL 540  
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAANPSPLETTKPDMLKEALKIAGFNEPENGNL 540  
Qy 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600  
Db 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600  
Qy 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Qy 661 DMLNISSLRODGKTFIDFKKYNKPLIISNPNYKNNYAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGKTFIDFKKYNKPLIISNPNYKNNYAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKIL---IFSKKG 731  
Db 721 IKKILKVVLGKKG 734

## RESULT 6

PCT-US94-01624-12  
; Sequence 12, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE AND CREW  
; STREET: Steuart Street Tower, 20th Floor, One Market  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PC-DOS/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01624  
; FILING DATE: June 25, 1993

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-115  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 903 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01624-12

Query Match 98.7%; Score 3722.5; DB 5; Length 903;  
Best Local Similarity 99.0%; Pred. No. 2.3e-253;  
Matches 727; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

Qy 1 EVKQENRLNESSSQGLLYFDFLNFAQPMVTSSTTGDLSIPSELENIPSENQYF 60

Db 1 EVKQENRLNESSSQGLLYFDFLNFAQPMVTSSTTGDLSIPSELENIPSENQYF 60  
Qy 61 QSAIWGSFIVKKSDEYTPATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWGSFIVKKSDEYTPATSDADNHVTWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKRKSTAGTSAGTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSSNSRKRKSTAGTSAGTVPDRN 180  
Qy 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNTHKKGITKYKSSPEKWTASDPYSDPEKVT 240  
Db 181 DGIPTDSLEVEGYVDVKNKRTFLSPWISNTHKKGITKYKSSPEKWTASDPYSDPEKVT 240  
Qy 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTONTSETRTTISKNTSTGRTH 300  
Db 241 GRIDKNVSPARPLVAAYPIVHVDMENILSKNEDQSTONTSETRTTISKNTSTGRTH 300  
Qy 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYNNVLPPTTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLPPTTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNNYQFLELEKTKQLRLDQVYGNIAVNFENGVRVDTGNNWSEV 480  
Db 421 LNAQKDFSSPTITMNNYQFLELEKTKQLRLDQVYGNIAVNFENGVRVDTGNNWSEV 480  
Qy 481 LPQIQETTARIIFNGKDLNLVERRIAAANPSPLETTKPDMLKEALKIAGFNEPENGNL 540  
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAANPSPLETTKPDMLKEALKIAGFNEPENGNL 540  
Qy 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600  
Db 541 QYQKDIETEDFNDOOTSONIKQLAELNATNIYVLDKIKLNKQMLIRDKRFHYDR 600  
Qy 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSGTGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Qy 661 DMLNISSLRODGKTFIDFKKYNKPLIISNPNYKNNYAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGKTFIDFKKYNKPLIISNPNYKNNYAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKIL---IFSKKG 731  
Db 721 IKKILKVVLGKKG 734

## RESULT 7

US-08-082-849B-31  
; Sequence 31, Application US/08082849B  
; Patent No. 5677274  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
; TITLE OF INVENTION: Related Methods  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/092,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-082-849B-31

Query Match 95.1%; Score 3590; DB 1; Length 719;  
Best Local Similarity 95.4%; Pred. No. 3.5e-244;  
Matches 705; Conservative 4; Mismatches 6; Indels 24; Gaps 2;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEYTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS---RKKRSTSGPTVP 176  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS---RKKRSTSGPTVP 180  
QY 177 DRNDGIPDSLEVEGYTVDNVKNKTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 236  
Db 181 DRNDGIPDSLEVEGYTVDNVKNKTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 240  
QY 237 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 296  
Db 241 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 300  
QY 297 RHTSEVEGNAEVAHSPFDIGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTAD 356  
Db 301 RHTSEVEGNAEVAHSPFDIGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTAD 360  
QY 357 TARLNANIRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKNL 416  
Db 361 TARLNANIRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKNL 420  
QY 417 APIALNAQKQFSPSTPIYMNQFLELEKTKQLRLDITDQVYGNATYNFENGRVVRDTGSN 476  
Db 421 APIALNAQKQFSPSTPIYMN-----YGNATYNFENGRVVRDTGSN 460  
QY 477 WSEVLPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMLKEALKIATGFNEP 536  
Db 461 WSEVLPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMLKEALKIATGFNEP 520  
QY 537 NGNLQYQKQKQITTEPFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNACNLIIRDKRF 596  
Db 521 NGNLQYQKQKQITTEPFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNACNLIIRDKRF 580  
QY 597 HYDRNNIAGVADESIVKAEHREVNSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 656  
Db 581 HYDRNNIAGVADESIVKAEHREVNSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 640

QY 657 NDYDMLNLISSLRQDGKTFIDFKKYNDKLPYISNPNYKVNVAVTKENTINPSENGDT 716  
Db 641 NDYDMLNLISSLRQDGKTFIDFKKYNDKLPYISNPNYKVNVAVTKENTINPSENGDT 700  
QY 717 STNGIKKILIPSKGYEIG 735  
Db 701 STNGIKKILIPSKGYEIG 719

RESULT 8  
PCT-US94-01624-31  
Sequence 31, Application PC/TUS9401624  
GENERAL INFORMATION:  
APPLICANT: Leppa, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: ANTHEAX TOXIN FUSION PROTEINS AND  
RELATED METHODS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW  
STREET: Steuart Street Tower, 20th Floor, One Market  
STREET: Plaza  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/01624  
FILING DATE: June 25, 1993  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-115  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 31:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01624-31

Query Match 95.1%; Score 3590; DB 5; Length 719;  
Best Local Similarity 95.4%; Pred. No. 3.5e-244;  
Matches 705; Conservative 4; Mismatches 6; Indels 24; Gaps 2;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEYTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNS---RKKRSTSGPTVP 176  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNTATIMQRGNFLOGPTVP 180  
QY 177 DRNDGIPDSLEVEGYTVDNVKNKTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 236  
Db 181 DRNDGIPDSLEVEGYTVDNVKNKTFATSDNHNVTMWDDQEVINKASNKNIRLEKGLYQIKIY 240  
QY 237 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 296  
Db 241 EKVTRIDKNVSPARPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTS 300  
QY 297 RHTSEVEGNAEVAHSPFDIGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTAD 356  
Db 301 RHTSEVEGNAEVAHSPFDIGGSVSAFNSNSSTVAIDHSLSLAGERTWAETMGLNTAD 360  
QY 357 TARLNANIRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKNL 416  
Db 361 TARLNANIRYNTGTATYVNLPTTSLVGLKNOTLATIKAKENQLSOILAPNNYPSKNL 420  
QY 417 APIALNAQKQFSPSTPIYMNQFLELEKTKQLRLDITDQVYGNATYNFENGRVVRDTGSN 476  
Db 421 APIALNAQKQFSPSTPIYMN-----YGNATYNFENGRVVRDTGSN 460  
QY 477 WSEVLPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMLKEALKIATGFNEP 536  
Db 461 WSEVLPOIQTETARIIFNGKDLNIVERRIAANVPSDPLETTKPDMLKEALKIATGFNEP 520  
QY 537 NGNLQYQKQKQITTEPFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNACNLIIRDKRF 596  
Db 521 NGNLQYQKQKQITTEPFDNFQDQTSQNIKNQLAELNATNIYVLDKIKLNACNLIIRDKRF 580  
QY 597 HYDRNNIAGVADESIVKAEHREVNSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 656  
Db 581 HYDRNNIAGVADESIVKAEHREVNSTEGLLNIDKIDKILSGYIVIEDTEGLKEVI 640

QY 237 EKVTGRIDKNVSPARHPPLVAAPIVHVDMENILISKNEDQSTONTDSETRTISKNTSTS 296  
DB 241 EKVTGRIDKNVSPARHPPLVAAPIVHVDMENILISKNEDQSTONTDSETRTISKNTSTS 300  
QY 297 RTHTSVHGNAEVAHSAFFDGGSVSAGFSNSNSTVAIDHSLSLAGERWTAETMGLNTAD 356  
DB 301 RTHTSVHGNAEVAHSAFFDGGSVSAGFSNSNSTVAIDHSLSLAGERWTAETMGLNTAD 360  
QY 357 TARLNANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENOLSOILLAPNNYPSKNL 416  
DB 361 TARLNANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENOLSOILLAPNNYPSKNL 420  
QY 417 APIALNAQDFSTPTIMNYNQFLEKTKQLRLDQDVYGNATYFENGVRVDTGSN 476  
DB 421 APIALNAQDFSTPTIMN-----YGNATYFENGVRVDTGSN 460  
QY 477 WSVLPQIOETTHARILFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIARFGENEP 536  
DB 461 WSVLPQIOETTHARILFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIARFGENEP 520  
QY 537 NGNLQYQGDITBFDNFQDOTSONIKNOAELNATNIYVLDKIKLAKMMILIRDKRF 596  
DB 521 NGNLQYQGDITBFDNFQDOTSONIKNOAELNATNIYVLDKIKLAKMMILIRDKRF 580  
QY 597 HYDRNNIYGADESUVKEAHREVINSGTEGLLNIDKIRKILSGYVIEIDTEGLKEVI 656  
DB 581 HYDRNNIYGADESUVKEAHREVINSGTEGLLNIDKIRKILSGYVIEIDTEGLKEVI 640  
QY 657 NDYDMNLISLQDQGTFTDFPKYNDKPLVYISPNYKNVYAVTKENTIINPSENGDT 716  
DB 641 NDYDMNLISLQDQGTFTDFPKYNDKPLVYISPNYKNVYAVTKENTIINPSENGDT 700  
QY 717 STNGIKKILFSGKGYEIG 735  
DB 701 STNGIKKILFSGKGYEIG 719

## RESULT 9

US-09-273-839A-8  
; Sequence 8, Application US/09273839A  
; Patent No. 6329156  
; GENERAL INFORMATION:  
; APPLICANT: Cirino, Nick M  
; APPLICANT: Jackson, Paul J  
; APPLICANT: Lehnert, Bruce E  
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface  
; FILE REFERENCE: S-89,662  
; CURRENT APPLICATION NUMBER: US/09/273,839A  
; CURRENT FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-273-839A-8

Query Match 33.8%; Score 1275; DB 4; Length 288;  
Best Local Similarity 100.0%; Pred. No. 5.2e-82;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 466 ETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIARFGENEPNGLQYQK 545  
DB 29 ETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIARFGENEPNGLQYQK 88  
QY 546 DITEFDNFQDOTSONIKNOAELNATNIYVLDKIKLAKMMILIRDKRFHYDRNNIAV 605  
DB 89 DITEFDNFQDOTSONIKNOAELNATNIYVLDKIKLAKMMILIRDKRFHYDRNNIAV 148  
QY 606 GADESUVKEAHREVINSGTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRYDMLNI 665  
DB 149 GADESUVKEAHREVINSGTEGLLNIDKIRKILSGYVIEIDTEGLKEVINDRYDMLNI 208

QY 666 SSLRQDQGTFTDFPKYNDKPLVYISPNYKNVYAVTKENTIINPSENGDTSTNGIKKIL 725  
DB 209 SSLRQDQGTFTDFPKYNDKPLVYISPNYKNVYAVTKENTIINPSENGDTSTNGIKKIL 268  
QY 726 IFSKKGYEIG 735  
DB 269 IFSKKGYEIG 278

## RESULT 10

US-08-960-780-32  
; Sequence 32, Application US/08960780  
; Patent No. 6204435  
; GENERAL INFORMATION:  
; APPLICANT: Feitelson, Jerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schmeits, James  
; APPLICANT: Loewer, David  
; APPLICANT: Dullum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide  
; TITLE OF INVENTION: Sequences Which Encode These Toxins  
; NUMBER OF SEQUENCES: 134  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32608-6669  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/960,780  
; FILING DATE: 30-OCT-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,848  
; FILING DATE: 30-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: MA-708  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 881 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: 177C8  
US-08-960-780-32

Query Match 20.7%; Score 781; DB 3; Length 881;  
Best Local Similarity 30.4%; Pred. No. 1.4e-46;  
Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

QY 4 QENLLNSESQGLGYFSDLNFOAPVMVTSSTGDLSPSELEN--IPSENOYFQ 61  
DB 39 QKQK---QKEMDRKGLGYFGKDF-SNLTMFAPTRDSTLIYDQOTANKLDDKQSEYQ 94  
QY 62 SAIMSGFVKVSKDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQYQ 121

Db 95 SIRWIGLQSKETGDTFNLSEDEQAIIIEINGKISNGKEKQVHLEKGLVPIKIEYQ 154  
Qy 122 RENPTKGLD-----FKLYWTDSONKKEVSSDNLQPLKQKSS-----N 162  
Db 155 SD--TKFNIDSKTPELKLFKIDSQVQVQOQDELNPFNFKEQEFLLAKPSKINLFT 212  
Qy 163 SRKKESTAGPTVPRDNDGIPDSLEVEGYTVVKNKETFPLSPWISNIHEKGLTKYKSS 222  
Db 213 QMKREIDE-----TDTGDSIPDLWEENGYTI-----QNRIVAKWDDSL-ASKGYTKFVSN 264  
Qy 223 PEKWTASDPYSDFKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNT 282  
Db 265 PLESHTVGDPTDYKAARDLDSNAKETFNPLVAAPPSVNVSMKEKVLSPENLS----- 320  
Qy 283 DSETRTSKNTSTSRTHSTSEVHGNAEVSFFDGGSGVSAGFSNSSS---TVAIDHSL 339  
Db 321 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQSHSET 363  
Qy 340 LAGERTWAETMG-----LNTADTARINANIRVYNTGTAPIYVNLPTTSLVLGKNOTLATIK 395  
Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANVRNNVGTGAIYVKPTTSFVL--NNDTIATIT 420  
Qy 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDFSTPTTMVYNQFLEKTKQLDLTDQV 455  
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNSHPITLNKKQVDNLLNKKPMLETNQT 480  
Qy 456 YGNIATYNENGRVAVDTGSNWSVLPQIQTETARIIFNGKDLNIVERRIAAVNSPDPLE 515  
Db 481 DG---VYKIKDTHGNIIVTGENWNGVIOQIKAKTASIIIVDDGE-RVAEKVAAKDYENPED 536  
Qy 516 TTKPDMTLKALIAF--GFNPNGLQYQGDITETDF--NFDQOTSQNTKNQLAEL-- 569  
Db 537 KT-PSLTLKDALKSLPYDEIKIEGLLYYKRPYESSVMTYLDENTAKETVKQJNDTTG 595  
Qy 570 ---NATNIYVLDKIKLNKAMILLRDKRFHYDRNNIAGVADSVVKAHREVINSSTEG 626  
Db 596 KFDVSHLYDV-----KLTFCAMNTIK-LSILYDN---AESDNSIGKWTNIVSGGNG 647  
Qy 627 -----LLANID-----KDIRKILSYIVIEIDTE-----GLKE 654  
Db 648 KQOYSSNNPDANLTNTDAQELKNRDYVVISLYMKSEKNTQCEITIDGIEIYPTTKTVN 707  
Qy 655 VINDRVDMLNI--SSLRQDKTFIDFKYNDKLPVVISPNKYKNVYVTKENTLINPSE 712  
Db 708 VNKDNYKRLDIIAHNKSPISIIHKT-NDEITLFWDDISI--TDVASIKPEN--LTDSE 763  
Qy 713 NGDT-STNGIK---KILIFSKGYEIG 735  
Db 764 IKQIYSRYGKLEDGILIDKKGGIHYG 790

## RESULT 11

US-09-073-898-32  
; Sequence 32, Application US/09073898  
; Patent No. 6242669  
; GENERAL INFORMATION:  
; APPLICANT: Feitelson, Jerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schneits, James  
; APPLICANT: Loewer, David  
; APPLICANT: Dullum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; APPLICANT: Morrill, George  
; APPLICANT: Finstad-Lee, Stacey  
; TITLE OF INVENTION: No. 6242669a1 Pesticidal Toxins and Nucleotide  
; TITLE OF INVENTION: Sequences Which Encode These Toxins  
; NUMBER OF SEQUENCES: 144  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/073,898  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,848  
FILING DATE: 30-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/960,780  
FILING DATE: 30-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-708C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 881 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: PSI177C8  
US-09-073-898-32

Query Match 20.7%; Score 781; DB 3; Length 881;

Best Local Similarity 30.4%; Pred. No. 1.4e-46;

Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

Qy 4 QENRLNESSSQGLLYGYFDLNFQAPMVTSTTGDLSTPSSLEN--TPSENQYFQ 61  
Db 39 QKQK--QKMDRKGILLYGYFKGKDF-SNLTMFAPTRDSTLIYDQQTANKLDDKKQOEYQ 94  
Qy 62 SAINSGIKYKKSDEYTFATSNHVTWVDDEVINKASNNKIRLEKRLYQIKIYQ 121  
Db 95 SIRWIGLQSKETGDTFNLSEDEQAIIIEINGKISNGKEKQVHLEKGLVPIKIEYQ 154  
Qy 122 RENPTKGLD-----FKLYWTDSONKKEVSSDNLQPLKQKSS-----N 162  
Db 155 SD--TKFNIDSKTPELKLFKIDSQVQVQOQDELNPFNFKEQEFLLAKPSKINLFT 212  
Qy 163 SRKKESTAGPTVPRDNDGIPDSLEVEGYTVVKNKETFPLSPWISNIHEKGLTKYKSS 222  
Db 213 QMKREIDE-----TDTGDSIPDLWEENGYTI-----QNRIVAKWDDSL-ASKGYTKFVSN 264  
Qy 223 PEKWTASDPYSDFKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNT 282  
Db 265 PLESHTVGDPTDYKAARDLDSNAKETFNPLVAAPPSVNVSMKEKVLSPENLS----- 320  
Qy 283 DSETRTSKNTSTSRTHSTSEVHGNAEVSFFDGGSGVSAGFSNSSS---TVAIDHSL 339  
Db 321 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQSHSET 363  
Qy 340 LAGERTWAETMG-----LNTADTARINANIRVYNTGTAPIYVNLPTTSLVLGKNOTLATIK 395  
Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANVRNNVGTGAIYVKPTTSFVL--NNDTIATIT 420  
Qy 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDFSTPTTMVYNQFLEKTKQLDLTDQV 455  
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNSHPITLNKKQVDNLLNKKPMLETNQT 480

```

QY 456 YGNIATYFNENGRVVDGTSNWSSEVLPOIQTETARIENGKDLNVERIAAVNPSPLE 515
Db 481 DG---VYKIDTHGNIVTGGWNGVIOQIKAKTASIIVDDGE-RVAEKVAAKDYENPED 536
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLOYQOGKITEFDF--NFDQOTSQNIKNQIAEL-- 569
Db 537 KT-PBLTLKDALKLSYDPEIKEIEGLLYKPKPIYESSVMYLDENTAKEVTKQLNDITG 595
QY 570 ---NATNIYTVLDKIKLNAKNILIRKFRHYDRNNAIAGADESVVKEAHEVINSSTEG 626
Db 596 KFKDVSHLYDV---KLTPKANVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647
QY 627 -----LLNID-----KDIRKILSGYIVIEIDTE-----GLKE 654
Db 648 KKQYSSNPNPDANLTNTDAQKLNKNRDYVYISLYMKSEKNTQCEITIDGEIYPIITKTWN 707
QY 655 VINDRYDMLNI--SSLRODGKTFIDFKKYNKDLPLIYISNPNYKVNVAATKENTIIINPSE 712
Db 708 VNKNYKRLDIIAHNIKSNPISSIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735
Db 764 IKQIYSRYGKLEDGLIDKGGIHVG 790

RESULT 12
US-09-307-106-8
; Sequence 8, Application US/09307106
; Patent No. 6603063
; GENERAL INFORMATION:
; APPLICANT: Fetteison, Jerald S.
; APPLICANT: Schnept, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/307,106
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/073,898
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355

```

```

; REFERENCE/DOCKET NUMBER: MA-708C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 881 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: PS177C8a
; US-09-307-106-8

```

```

Query Match 20.7%; Score 781; DB 4; Length 881;
Best Local Similarity 30.4%; Pred. No. 1.4e-46;
Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

QY 4 QENRLNESSSQGLGYFSDLNQAPMVTTSITGDLSPSSLEN--IPSENQYFQ 61
Db 39 QKQ---QKEMDRKGLLGYFKGDF-SNLTWFAPTDSTLIYDQOTANKLLKKQEQY 94
QY 62 SAIWSGFIVKVKDEYTFATSDAHNVMTVMVDDEVINKASNSKIRLEKRLYQIKIYQ 121
Db 95 SIRWIGLIQSKETGDTFNLSDEQALIEINGKIKSNKKEKQVHLEKGLVPIKIEYQ 154
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISSDNLQLPKOKSS-----N 162
Db 155 SD---TKFNIDSKTFKELKFKIDSQSQVQVQODELENPEFNKESQEFLLAKSKINLFT 212
QY 163 SRKRSTSGAGTVPDRDNDGIPDSLEVEGYTVDKNKRFTLSPWISNIHEKGLTKYKSS 222
Db 213 QKMKREIDED---TDTGDSIPDLWEENGYTI---QNRIAVKWDDSL-ASKGYTKFVSN 264
QY 223 PEKWTASDPYSDEKVTGRIDKNVSPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNT 282
Db 265 PLESHTVGDPTDYDKAARDLDSNAKETPNLVAAPPVSVNMEKVLSPNENLS----- 320
QY 283 DSETRISKNTSTSRHTSEVHGNAEVHASFFDIGSVSAGFSNSNS-----TVAIDHSL 339
Db 321 ---NSVESHSSTWVSYT-----NTE-----GASVEAGIGPKGISFGVSNYQHS 363
QY 340 LAGRTWAETWG-----INTADTARLNANIRYVNTGTAPIYVNLPTSLVLGKQCTLATIK 395
Db 364 VAQE--WGTSTGNTSQFTASAGYLNANVRNNVGTGAIYDVKPTTSFVL-NNDTIIAT 420
QY 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDFSTPTIMYNOFLEKTKQLRLDQV 455
Db 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHPITLKKQVDNLLNKKPMLETNQT 480
QY 456 YGNIATYFNENGRVVDGTSNWSSEVLPOIQTETARIENGKDLNVERIAAVNPSPLE 515
Db 481 DG---VYKIDTHGNIVTGGWNGVIOQIKAKTASIIVDDGE-RVAEKVAAKDYENPED 536
QY 516 TTKPDMTLKEALKIAF--GFNEPNGNLOYQOGKITEFDF--NFDQOTSQNIKNQIAEL-- 569
Db 537 KT-PBLTLKDALKLSYDPEIKEIEGLLYKPKPIYESSVMYLDENTAKEVTKQLNDITG 595
QY 570 ---NATNIYTVLDKIKLNAKNILIRKFRHYDRNNAIAGADESVVKEAHEVINSSTEG 626
Db 596 KFKDVSHLYDV---KLTPKANVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647
QY 627 -----LLNID-----KDIRKILSGYIVIEIDTE-----GLKE 654
Db 648 KKQYSSNPNPDANLTNTDAQKLNKNRDYVYISLYMKSEKNTQCEITIDGEIYPIITKTWN 707
QY 655 VINDRYDMLNI--SSLRODGKTFIDFKKYNKDLPLIYISNPNYKVNVAATKENTIIINPSE 712
Db 708 VNKNYKRLDIIAHNIKSNPISSIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763
QY 713 NGDT-STNGIK---KILIFSKGYEIG 735

```

Db 764 IKQYRYGKLEGGILIDKGGIHYG 790

RESULT 13

US-09-850-351A-32

Sequence 32, Application US/09850351A

Patent No. 6656908

GENERAL INFORMATION:

APPLICANT: Feltelson, Gerald S.

Schnepi, H. Ernest

Narva, Kenneth B.

Stockhoff, Brian A.

Schmeits, James

Loewer, David

Dullum, Charles Joseph

Muller-Cohn, Judy

Stamp, Lisa

Morrill, George

TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide Sequences Which Encode These Toxins

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606-6669

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/850,351A

FILING DATE: 07-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/073,898

FILING DATE: 06-MAY-1998

APPLICATION NUMBER: US 08/960,780

FILING DATE: 30-OCT-1997

APPLICATION NUMBER: US 60/029,848

FILING DATE: 30-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708CD1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 881 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: PS177C8

SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-09-850-351A-32

Query Match 20.7%; Score 781; DB 4; Length 881;

Best Local Similarity 30.4%; Pred. No. 1.4e-46;

Matches 245; Conservative 134; Mismatches 298; Indels 130; Gaps 34;

Qy 4 QENRLNSESQGLIGYFSLNFQAPMVVTSSTGDLSPSSLEN--IPSENQYFQ 61

Db 39 QKQKQ--QKEMDRKGLLYGFKGDF-SNLTMPAPTRDSTLIYDQQTANKLLDKKQEQY 94

Qy 62 SATWSGFIKVKSDVEYFATSDAHVMTWYDDQEVINKASNSKNKLEKGRLYQIKIQY 121

Db 95 SIRWIGLIQSKETGDTFTNLSEDOAIIEINGKIISNKKKEQVHVLEKGLVPIKIEYQ 154

Qy 122 RENPTEKGLD-----FKLYWTDSONKKEVIESDNLQJPELKQKSS-----N 162

Db 155 SD--TKFNIDSKTFKELKLFKIDSQNQPOQOQDELNPEFNKKESQEBFLAKPSKINLFT 212

Qy 163 SRKRSSTAGPTVDRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIEHKKGLTKYKSS 222

Db 213 QKKREIDED---TDTGDSIPDLWEENGYTI-----QNRPIAVKWDSDL-ASKGYTKFVSN 264

Qy 223 PEKWTASDPYSDPFKVTGRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDOSTQNT 282

Db 265 PLESHTVGDPYTDYEKAARDLDSNAKETFNELVAAFPSVNVSMKVLISPNENLS---- 320

Qy 283 DSETRISKNTSTSRTHSEVHGNAEVHASFFDIGGSVAGFSNNS--TVADHSLS 339

Db 321 ----NSVESHSINWSYT-----NTE-----GASVEAGIGPKGISFGVSNYVCHSET 363

Qy 340 LAGERTWAETMG---LNTADTARLNANIRYVNTGPAPYVNLPTTSLVLGKNQTLATIK 395

Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANRYNNVGTGAIYDVKPTTSFVL--NNDTIATIT 420

Qy 396 AKENQLSQILAPNNYPSKNLAPIALNAQKDSSTBITWNYQFLEKTKOLRLDIDOV 455

Db 421 AKSNSTALNISPGESYFKGQNGIATSMDDFNHSBITLKKQVDNLLNKKPMLENTQT 480

Qy 456 YGNIATYFNGRVRVDTGSNMSEVLPIQIETTARIIFNGKDLNLVERIAAVNPSDPLE 515

Db 481 DG---VYKIKDTRGNIVTGGWNGVIQIKAKTASIIVDDGE--RVAEKVAAKDYENPED 536

Qy 516 TYPDMTLKEALKIAP--GFNEPENGLOYGKDIETEDF--NPDQOTSONIKNQLAEL-- 569

Db 537 KT-PSLTLKDALKLSYDPEIKETEGLLYKPKFYESSWVTLDENTAKEVTKQLNDTTG 595

Qy 570 ---NATNIYTVLDKIKLNAMNIIIRDGRPHYDRNNIAGVADSVVKEAHRVINSSTEG 626

Db 596 KFKDVSHLYDV---XLTPTQNVTIK-LSILYDN---AESNDNSIGKWTNTIVSGGNG 647

Qy 627 -----LLNID-----KDIRKLSGYIYEIEDTE-----GLXE 654

Db 648 KKOYSSNNPDANLTANTDAQEKLNKRDYIISLYMKSEKNTCEITIDGIEIYPIITKTVN 707

Qy 655 VINDRYDMLNI--SSLRQDGKTFIDFKYNDKLPYISNPYKVNKYAVTKENTIINPSE 712

Db 708 VKNKYKELDIIAHNIKSNFISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763

Qy 713 NGDT-STNGIK---KILIFSKKGYEIG 735

Db 764 IKQYRYGKLEGGILIDKGGIHYG 790

RESULT 14

US-08-471-033-5

Sequence 5, Application US/08471033

Patent No. 5770696

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Koziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: Carr, Brian

APPLICANT: Desai, Nalini M

APPLICANT: Kostichka, N. Kristy

APPLICANT: Duck, Nicholas B

APPLICANT: Estruch, Juan J

TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

QY 516 TTKPDMTLKALFIAP--GFNEPNGNLQYQKIDITEFDF--NPDQOTSONIKNQLAEL-- 569  
Db 540 KT-PSITLKDALKLSYDEIKIEGLLYKKNPIYESSVMTYLDENTAKEVTQLNDTTG 598  
QY 570 ---NATNIYTVLDKIKLNAMKILIRDKRPHDRNNAIAGADESVVKEAHRVINSSTEG 626  
Db 599 KFZDVSHLYDV---KLTTPKMNVTIK-LJSLYDN---AESNDNSIGKNTNTNIVSGNGK 650  
QY 627 -----LLNID-----KDIRKILSGYIIVEIDTE-----GLKE 654  
Db 651 KKOYSSNPNPANLTNTDAQELKLNKRDYIISLYMKSEKNTQCEITIDGIEIPIITKTVN 710  
QY 655 VINDRYMDLN-----ISSLRDQGTFTDFKFKYNDKLPYISNPNYKVNVAVTKE 704  
Db 711 VNKDNYKRLDIIAHNIKSNPISLH-----IKTNDIEITLFWDDISI-TDVASIKPE 760  
QY 705 NTIINPSENGDT-STNGIK---KILIFSKGYEIG 735  
Db 761 N-LTDSEIKQYISRYGKLEDCILIDKGGIHYG 793

RESULT 15  
US-08-471-044-5  
; Sequence 5, Application US/08471044  
; Patent No. 5840868  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Gregory W  
; APPLICANT: Koziel, Michael G  
; APPLICANT: Mullins, Martha A  
; APPLICANT: Nye, Gordon J  
; APPLICANT: Carr, Brian  
; APPLICANT: Desai, Nalini M  
; APPLICANT: Kostichka, N. Kristy  
; APPLICANT: Duck, Nicholas B  
; APPLICANT: Estruch, Juan J  
; TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30B  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,044  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/463,483  
; FILING DATE: 05-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/314,594  
; FILING DATE: 09-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/218,018  
; FILING DATE: 23-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/037,057  
; FILING DATE: 25-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pace, Gary M.  
; REGISTRATION NUMBER: 40,403  
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SOLV3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-541-8582  
; TELEFAX: 919-541-8689

QY 516 TTKPDMTLKALFIAP--GFNEPNGNLQYQKIDITEFDF--NPDQOTSONIKNQLAEL-- 569  
Db 540 KT-PSITLKDALKLSYDEIKIEGLLYKKNPIYESSVMTYLDENTAKEVTQLNDTTG 598  
QY 570 ---NATNIYTVLDKIKLNAMKILIRDKRPHDRNNAIAGADESVVKEAHRVINSSTEG 626  
Db 599 KFZDVSHLYDV---KLTTPKMNVTIK-LJSLYDN---AESNDNSIGKNTNTNIVSGNGK 650  
QY 627 -----LLNID-----KDIRKILSGYIIVEIDTE-----GLKE 654  
Db 651 KKOYSSNPNPANLTNTDAQELKLNKRDYIISLYMKSEKNTQCEITIDGIEIPIITKTVN 710  
QY 655 VINDRYMDLN-----ISSLRDQGTFTDFKFKYNDKLPYISNPNYKVNVAVTKE 704  
Db 711 VNKDNYKRLDIIAHNIKSNPISLH-----IKTNDIEITLFWDDISI-TDVASIKPE 760  
QY 705 NTIINPSENGDT-STNGIK---KILIFSKGYEIG 735  
Db 761 N-LTDSEIKQYISRYGKLEDCILIDKGGIHYG 793

Query Match 20.6%; Score 779; DB 1; Length 884;  
Best Local Similarity 30.3%; Pred. No. 1.9e-46;  
Matches 247; Conservative 129; Mismatches 293; Indels 146; Gaps 34;  
QY 4 QENRLNSESSESSQGLLYGYFDLNFQAPMVVTSSTTGDLSIPSELEN--IPSENQYFQ 61  
Db 42 QXNQ---QKEMRKGLLYGYFKGDF-SNLTWFAETRDSTLIYDOOTANKLLDKKQEQYQ 97  
QY 62 SAIWGFIKVKSDVETPATSADNHVTMMVDQEVINKASNSKTRLEKGLYQIKIOYQ 121  
Db 98 SIRWIGLIQSKETGDTFNLSEDEQAIIIEINGKIISNKGKQVHVHLEKGLVPKIEYQ 157  
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVISSDNLQLPELKQSS-----N 162  
Db 158 SD--TKFNIDSTFKELKLFKLDSONQOQOQDELRNPEFNKESQEFKAKPSKINLFT 215  
QY 163 SRKGRSTASGPTVPRDNDGIPDSLEVEGYTVVKNKRTFLSPWTSNIHEKGLTKYKSS 222  
Db 216 QRMKREIDED---TDTGDSIPDLWEENGYTI---QNRIAVKWDDSL-ASKGYTKFYSN 267  
QY 223 PEKMGSTASDPYDFEKTGRIDKNVSPEARHPLVAAPYPIVHVDMENIILSKNEDOSTQNT 282  
Db 268 PLESHTVGPDIYDKAARDLDSNAKETFNELVAAPFVSVNVMSEKVLSPNENLS---- 323  
QY 283 DSETRTISKNTSTSTHTSEVHGNAEVHASFEDIGGSVSAFNSNSS---TVAIDHSL 339  
Db 324 ---NSVESHSSTNYSY---NTE-----GASVEAGIGPKGISFGVSVNYQHSET 366  
QY 340 LAGERTWATMG-----LNTADTARINANRYNTGTAPYVLPPTSVLGKQNTLAIK 395  
Db 367 VAGE--WGTSNGTSQFNASAGYLNANRYNVGTGAIYDVKPTTSFVL--NNDTIAT 423  
QY 396 AKENQLSQILAPNNYPSKNLAPIALNAQDSSPTITMNNYQFLEKTKQLRLDTPQV 455  
Db 424 AKSNSTALNISPGESYPKKGQNGIAITSDWDNSHPITLTKKQVNDLLNNKPMLEMTQ 483  
QY 456 YGNIATYFNENGRVVDTSNISEVLPOIQTETARIIFNGKDLNLVERRIAAVNSPDL 515  
Db 484 DG---VYKIDTHGNIVTGEWNGVIOQIKAKTASTIIVDDGS-RVAEKRAVKADYENPED 539



```

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 884 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-471-044-5

```

Search completed: May 3, 2004, 19:42:40  
Job time : 18.0986 secs



Db 61 QSAIWSGFIKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPENGL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPENGL 540  
QY 541 QYQKGLITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLIRDKRPHYDR 600  
Db 541 QYQKGLITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLIRDKRPHYDR 600  
QY 601 NNTAVGADESUVKEAREVINSSTEGLLNIDKIDKIRKLSGYIVEEDTEGLKEVINDRY 660  
Db 601 NNTAVGADESUVKEAREVINSSTEGLLNIDKIDKIRKLSGYIVEEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735

RESULT 2

US-09-848-909-1  
; Sequence 1, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-1

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLINESSSQGLLYYFSDLNFCQPMVVTSTTGDLSIPSSSELENIIPSENOYF 60  
Db 1 EVKQENRLINESSSQGLLYYFSDLNFCQPMVVTSTTGDLSIPSSSELENIIPSENOYF 60  
QY 61 QSAIWSGFIKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120  
Db 61 QSAIWSGFIKVKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLOLPELKOKSNSRKRSTSGPTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKSTASDPYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSTRITISKNSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGRVVDTSNWSSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPENGL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNPENGL 540  
QY 541 QYQKGLITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLIRDKRPHYDR 600  
Db 541 QYQKGLITEFDNFDDQTSQNIKNQLAELNATNIYVLDKIKLNAKONLIRDKRPHYDR 600  
QY 601 NNTAVGADESUVKEAREVINSSTEGLLNIDKIDKIRKLSGYIVEEDTEGLKEVINDRY 660  
Db 601 NNTAVGADESUVKEAREVINSSTEGLLNIDKIDKIRKLSGYIVEEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735

RESULT 3

US-09-848-909-2  
; Sequence 2, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis

US-09-848-909-2

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
QY 61 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNLOLPELKOKSSNSRKRSTSGAGTVPPDRN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNLOLPELKOKSSNSRKRSTSGAGTVPPDRN 180  
QY 181 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 181 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTIMYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTIMYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
QY 541 QYQKDIITFDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRFHYDR 600  
Db 541 QYQKDIITFDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIETEGLEKVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIETEGLEKVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

RESULT 4  
US-09-848-909-3  
; Sequence 3, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, Brett R.  
; APPLICANT: Collier, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04

US-09-848-909-3

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIIPSENQYF 60  
QY 61 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNLOLPELKOKSSNSRKRSTSGAGTVPPDRN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVIVSSDNLOLPELKOKSSNSRKRSTSGAGTVPPDRN 180  
QY 181 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 181 QSAIWGSGFIKVKKSDRYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTIMYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTIMYNOFLEKTKQLRLDQVYGNIAIYFNGRVRVDTGSNWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
QY 541 QYQKDIITFDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRFHYDR 600  
Db 541 QYQKDIITFDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKQNTLIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIETEGLEKVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIETEGLEKVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

RESULT 5  
US-09-848-909-4  
; Sequence 4, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-4

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVTSSITGDLSPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVTSSITGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKY 120  
DB 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKY 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
QY 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGKLTYSKSPKWSASDPSDFEKT 240  
DB 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGKLTYSKSPKWSASDPSDFEKT 240  
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480  
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPENG 540  
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPENG 540  
QY 541 QYQKDIETEDFNFDOQTSONIKQLAELNATNIYTVLDKIKNAKNILLIRKRFHYDR 600  
DB 541 QYQKDIETEDFNFDOQTSONIKQLAELNATNIYTVLDKIKNAKNILLIRKRFHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
DB 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNISLRQDKTFFIDFKKYNKDLPLYISNPNYKNVAVTKENTIIINPSNGDTSTNG 720  
DB 661 DMLNISLRQDKTFFIDFKKYNKDLPLYISNPNYKNVAVTKENTIIINPSNGDTSTNG 720  
QY 721 IKKILIPSKKGYEIG 735  
DB 721 IKKILIPSKKGYEIG 735

RESULT 6  
US-09-848-909-5

; Sequence 5, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-5

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVTSSITGDLSPSSSELENIPSENQYF 60  
DB 1 EVKQENRLNESSSQGLLYFSDLNFCAPMVTSSITGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKY 120  
DB 61 QSAIWSGFIKVKSDSYTFATSDADNHVTMWVDDQEVINKASNKIRLEKGRLYQIKY 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
QY 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGKLTYSKSPKWSASDPSDFEKT 240  
DB 181 DGIPTDSEVEGYVDVKNKRTFLSPWISNIHEKKGKLTYSKSPKWSASDPSDFEKT 240  
QY 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480  
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFGVRVDTGNSWSEV 480  
QY 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPENG 540  
DB 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPENG 540  
QY 541 QYQKDIETEDFNFDOQTSONIKQLAELNATNIYTVLDKIKNAKNILLIRKRFHYDR 600  
DB 541 QYQKDIETEDFNFDOQTSONIKQLAELNATNIYTVLDKIKNAKNILLIRKRFHYDR 600  
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
DB 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
QY 661 DMLNISLRQDKTFFIDFKKYNKDLPLYISNPNYKNVAVTKENTIIINPSNGDTSTNG 720  
DB 661 DMLNISLRQDKTFFIDFKKYNKDLPLYISNPNYKNVAVTKENTIIINPSNGDTSTNG 720  
QY 721 IKKILIPSKKGYEIG 735  
DB 721 IKKILIPSKKGYEIG 735

Db 721 IKKILIFSKGYEIG 735  
|||||  
RESULT 7  
US-09-848-909-6  
; Sequence 6, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-6  
Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120  
Db 61 QSAIWGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120  
Qy 121 QRENPTKGLDFKLYWTDSONKEVTSNQLQPELKQKSSNRKRSTAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKEVTSNQLQPELKQKSSNRKRSTAGTVPDRDN 180  
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240  
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVLPPTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDQVYGNATYVNFENGRVRVDTGNSWSEV 480  
Db 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDQVYGNATYVNFENGRVRVDTGNSWSEV 480  
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540  
Qy 541 QYQKDIITFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKLNKAKNILLIRDKRPHYDR 600  
Db 541 QYQKDIITFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKLNKAKNILLIRDKRPHYDR 600  
Qy 601 NNIAVGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVKAEHREVINSTEGLLNIDKIRKILSGYIVETEDTEGLKEVINDRY 660

Qy 661 DMLNISLRDGGTPTIDFKKYNKDLPLYISNPNYKVNVTYKENTIIINPSENGDSTNG 720  
Db 661 DMLNISLRDGGTPTIDFKKYNKDLPLYISNPNYKVNVTYKENTIIINPSENGDSTNG 720  
Qy 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735  
RESULT 8  
US-09-848-909-7  
; Sequence 7, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-7  
Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSQGLLYGFFDLNFOAPMVVTSSTTGGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120  
Db 61 QSAIWGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKYQ 120  
Qy 121 QRENPTKGLDFKLYWTDSONKEVTSNQLQPELKQKSSNRKRSTAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKEVTSNQLQPELKQKSSNRKRSTAGTVPDRDN 180  
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASDPYSDFEKT 240  
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVLPPTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVLPPTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDQVYGNATYVNFENGRVRVDTGNSWSEV 480  
Db 421 LNAQKDFSSPTIMVYVNFQLEKTKQLRLDQVYGNATYVNFENGRVRVDTGNSWSEV 480  
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPENG 540  
Qy 541 QYQKDIITFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKLNKAKNILLIRDKRPHYDR 600

Db 541 QYQKDIETDFNFDOQTSONIKNQLAELNATNIYTVLKDILKNAKNNILIRDKRPHYDR 600  
QY 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNLSLRQDQKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNLSLRQDQKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735  
RESULT 9  
US-09-848-909-8  
; Sequence 8, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-8

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATSDAHNVMTWVDDQEVINKASNNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEYTFATSDAHNVMTWVDDQEVINKASNNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNRKRSTSGAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNRKRSTSGAGTPVDRDN 180  
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSSTADDPYSDFEKVT 240  
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSSTADDPYSDFEKVT 240  
QY 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNNAEVHASFDDIGSVSAGFSNNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360  
Db 301 SEVHGNNAEVHASFDDIGSVSAGFSNNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLPVTLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLPVTLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSGSTPTMNYNFOFLEKTKQLRLDQVYGNIAATYNPENGRVDTGSGNWSEV 480  
Db 421 LNAQKDFSGSTPTMNYNFOFLEKTKQLRLDQVYGNIAATYNPENGRVDTGSGNWSEV 480

QY 481 LPOIQTETARIIPNGKDLNLVERRIAAVNPSPDLETTKPDMLKEALKIAFGNEPENG 540  
Db 481 LPOIQTETARIIPNGKDLNLVERRIAAVNPSPDLETTKPDMLKEALKIAFGNEPENG 540  
QY 541 QYQKDIETDFNFDOQTSONIKNQLAELNATNIYTVLKDILKNAKNNILIRDKRPHYDR 600  
Db 541 QYQKDIETDFNFDOQTSONIKNQLAELNATNIYTVLKDILKNAKNNILIRDKRPHYDR 600  
QY 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVEKAEHREVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660  
QY 661 DMLNLSLRQDQKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNLSLRQDQKTFIDFKKYNKDLPLYISNPNYKNNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735  
RESULT 10  
US-09-848-909-9  
; Sequence 9, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-9  
Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEYTFATSDAHNVMTWVDDQEVINKASNNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEYTFATSDAHNVMTWVDDQEVINKASNNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNRKRSTSGAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNNRKRSTSGAGTPVDRDN 180  
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSSTADDPYSDFEKVT 240  
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKWSSTADDPYSDFEKVT 240  
QY 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPARHPVLAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNNAEVHASFDDIGSVSAGFSNNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360  
Db 301 SEVHGNNAEVHASFDDIGSVSAGFSNNSSTVAIDHSLSLAGERTWAETMGLTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLPVTLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLPVTLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420

```
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTIMYVNOFLEKTKQLBLDDQVYGNATYVNFGRVVDVDTGNSWSEV 480
Db 421 LNAQKDFSSPTIMYVNOFLEKTKQLBLDDQVYGNATYVNFGRVVDVDTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNNAKONILIRDKRPHYDR 600
Db 541 QYQKDIITFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNNAKONILIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735
```

## RESULT 11

```
US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 736
; TYPE: PRN
; ORGANISM: Bacillus anthracis
US-09-848-909-10
```

```
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Db 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Qy 181 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTADDPYDFEKT 240
Db 181 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTADDPYDFEKT 240
Qy 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300
```

```
Qy 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARL 360
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTIMYVNOFLEKTKQLRLDDQVYGNATYVNFGRVVDVDTGNSWSEV 480
Db 421 LNAQKDFSSPTIMYVNOFLEKTKQLRLDDQVYGNATYVNFGRVVDVDTGNSWSEV 480
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPNGNL 540
Qy 541 QYQKDIITFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNNAKONILIRDKRPHYDR 600
Db 541 QYQKDIITFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKNNAKONILIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Db 661 DMLNISLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSNGDTSTNG 720
Qy 721 IKKILIPSKKGYEIG 735
Db 721 IKKILIPSKKGYEIG 735
```

## RESULT 12

```
US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 736
; TYPE: PRN
; ORGANISM: Bacillus anthracis
US-09-848-909-11
```

```
Query Match 99.8%; Score 3767; DB 12; Length 736;
Best Local Similarity 99.9%; Pred. No. 4.5e-274;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Db 61 QSAIWSGFIKVKSGDEYTFATSAONHVMTWVDQEVINKASNNKIRLEKRLYQIKIQY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNQLQPELKQKSSNRKRSSTAGTVPDRN 180
Qy 181 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTADDPYDFEKT 240
Db 181 DGPIDPSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTADDPYDFEKT 240
```



Db 181 DGIPLSLEVEGYTVDKNGRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEYV 240  
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Qy 301 SEVHGNAEVAHAFDIFGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHAFDIFGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQDDFSPTPTMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480  
Db 421 LNAQDDFSPTPTMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGNEPENG 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGNEPENG 540  
Qy 541 QYCGKDIETDFNFDOOTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRPHYDR 600  
Db 541 QYCGKDIETDFNFDOOTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRPHYDR 600  
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Qy 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKILFSKGYEIG 735  
Db 721 IKKILFSKGYEIG 735

RESULT 13  
US-09-848-909-12  
; Sequence 12, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-12

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQPMVVTSTTGDLSPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQPMVVTSTTGDLSPSSSELENIPSENOYF 60  
Qy 61 QSAIWSGFIVKKSDEYTFATSDNHVTMVDDEVINKASNNKIRLEKGRLYQIKIQY 120  
Db 61 QSAIWSGFIVKKSDEYTFATSDNHVTMVDDEVINKASNNKIRLEKGRLYQIKIQY 120

Qy 121 QRENTEKGLDFKLWYTDSONKKEVISSDNLQLPKQKSSNSRKRKSTASGTPVDRDN 180  
Db 121 QRENTEKGLDFKLWYTDSONKKEVISSDNLQLPKQKSSNSRKRKSTASGTPVDRDN 180  
Qy 181 DGIPLSLEVEGYTVDKNGRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEYV 240  
Db 181 DGIPLSLEVEGYTVDKNGRTFLSPWISNIHEKGLTKYKSPKXWSTASDPYDFEYV 240  
Qy 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Qy 301 SEVHGNAEVAHAFDIFGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHAFDIFGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQDDFSPTPTMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480  
Db 421 LNAQDDFSPTPTMNYNQFLEKTKQLRLDQVYGNATYVNFNGRVRVDTGNSWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGNEPENG 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGNEPENG 540  
Qy 541 QYCGKDIETDFNFDOOTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRPHYDR 600  
Db 541 QYCGKDIETDFNFDOOTSQNIKNQLAELNATNITVLDKIKLNKXNMLIRDKRPHYDR 600  
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Qy 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODGTFFIDFKYNDKPLYSNPNKYVAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKILFSKGYEIG 735  
Db 721 IKKILFSKGYEIG 735

RESULT 14  
US-09-848-909-13  
; Sequence 13, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-13

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred. No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQPMVVTSTTGDLSPSSSELENIPSENOYF 60

Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRN 180  
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASGPTVPDRN 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASGPTVPDRN 240  
Qy 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
Db 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
Qy 301 SEVHGNAEVSASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEITMGLNTADTARL 360  
Db 301 SEVHGNAEVSASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEITMGLNTADTARL 360  
Qy 361 NANIRVYNTGTAPIYNNVLTPTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYNNVLTPTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVYDVTGSNWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVYDVTGSNWSEV 480  
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
Qy 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNITVLDKIKLNKKNILIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNITVLDKIKLNKKNILIRDKRPHYDR 600  
Qy 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKNVYAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKNVYAVTKENTIIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

RESULT 15  
US-09-848-909-14  
; Sequence 14, Application US/09848909  
; Publication No. US20020039588A1  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909-14

Query Match 99.8%; Score 3767; DB 12; Length 736;  
Best Local Similarity 99.9%; Pred No. 4.5e-274;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFPQPMVVTSTTGDLSIPSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDQNKKEVYSSDNQLQLPELKQSSNSRKKSTASGPTVPDRN 180  
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASGPTVPDRN 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGKLTYYKSSPKKSTASGPTVPDRN 240  
Qy 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
Db 241 GRIDKNVSPKARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300  
Qy 301 SEVHGNAEVSASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEITMGLNTADTARL 360  
Db 301 SEVHGNAEVSASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEITMGLNTADTARL 360  
Qy 361 NANIRVYNTGTAPIYNNVLTPTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYNNVLTPTSLVLGKQNTLATIKAKENQLSQILLAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVYDVTGSNWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYNFENGVRVYDVTGSNWSEV 480  
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPENG 540  
Qy 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNITVLDKIKLNKKNILIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNITVLDKIKLNKKNILIRDKRPHYDR 600  
Qy 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHEVINSSTGLLNIDKIRKILSGYIIEIDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKNVYAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODQKTFIDFKKYNKDLPLYSNPNYKNVYAVTKENTIIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

Search completed: May 3, 2004, 20:02:04  
Job time : 53.6393 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds  
(without alignments)  
4086.665 Million cell updates/sec

Title: US-09-848-909A-8  
Perfect score: 3773  
Sequence: 1 EVKQENLLNESSSQGLL.....TSTGIKKILIFSKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1:	/cgn2_6/prodata/2/paa/PCTUS_ COMB.pcp.*	Sequence 8, Appli
2:	/cgn2_6/prodata/2/paa/US06_ COMB.pcp.*	Sequence 8, Appli
3:	/cgn2_6/prodata/2/paa/US07_ COMB.pcp.*	Sequence 7, Appli
4:	/cgn2_6/prodata/2/paa/US08_ COMB.pcp.*	Sequence 7, Appli
5:	/cgn2_6/prodata/2/paa/US09_ COMB.pcp.*	Sequence 6, Appli
6:	/cgn2_6/prodata/2/paa/US10_ COMB.pcp.*	Sequence 6, Appli
7:	/cgn2_6/prodata/2/paa/US11_ COMB.pcp.*	Sequence 5, Appli
8:	/cgn2_6/prodata/2/paa/US12_ COMB.pcp.*	Sequence 5, Appli
9:	/cgn2_6/prodata/2/paa/US13_ COMB.pcp.*	Sequence 10, Appli
10:	/cgn2_6/prodata/2/paa/US14_ COMB.pcp.*	Sequence 10, Appli
11:	/cgn2_6/prodata/2/paa/US15_ COMB.pcp.*	Sequence 20, Appli
12:	/cgn2_6/prodata/2/paa/US16_ COMB.pcp.*	Sequence 20, Appli
13:	/cgn2_6/prodata/2/paa/US17_ COMB.pcp.*	Sequence 21, Appli
14:	/cgn2_6/prodata/2/paa/US18_ COMB.pcp.*	Sequence 21, Appli
15:	/cgn2_6/prodata/2/paa/US19_ COMB.pcp.*	Sequence 30, Appli
16:	/cgn2_6/prodata/2/paa/US20_ COMB.pcp.*	Sequence 1, Appli
17:	/cgn2_6/prodata/2/paa/US21_ COMB.pcp.*	Sequence 1, Appli
18:	/cgn2_6/prodata/2/paa/US22_ COMB.pcp.*	Sequence 3, Appli
19:	/cgn2_6/prodata/2/paa/US23_ COMB.pcp.*	Sequence 4, Appli
20:	/cgn2_6/prodata/2/paa/US24_ COMB.pcp.*	Sequence 5, Appli
21:	/cgn2_6/prodata/2/paa/US25_ COMB.pcp.*	Sequence 5, Appli
22:	/cgn2_6/prodata/2/paa/US26_ COMB.pcp.*	Sequence 7, Appli
23:	/cgn2_6/prodata/2/paa/US27_ COMB.pcp.*	Sequence 8, Appli
24:	/cgn2_6/prodata/2/paa/US28_ COMB.pcp.*	Sequence 9, Appli
25:	/cgn2_6/prodata/2/paa/US29_ COMB.pcp.*	Sequence 11, Appli
26:	/cgn2_6/prodata/2/paa/US30_ COMB.pcp.*	Sequence 11, Appli
27:	/cgn2_6/prodata/2/paa/US31_ COMB.pcp.*	Sequence 12, Appli
28:	/cgn2_6/prodata/2/paa/US32_ COMB.pcp.*	Sequence 12, Appli
29:	/cgn2_6/prodata/2/paa/US33_ COMB.pcp.*	Sequence 13, Appli
30:	/cgn2_6/prodata/2/paa/US34_ COMB.pcp.*	Sequence 14, Appli
31:	/cgn2_6/prodata/2/paa/US35_ COMB.pcp.*	Sequence 15, Appli
32:	/cgn2_6/prodata/2/paa/US36_ COMB.pcp.*	Sequence 16, Appli
33:	/cgn2_6/prodata/2/paa/US37_ COMB.pcp.*	Sequence 17, Appli
34:	/cgn2_6/prodata/2/paa/US38_ COMB.pcp.*	Sequence 18, Appli
35:	/cgn2_6/prodata/2/paa/US39_ COMB.pcp.*	Sequence 19, Appli
36:	/cgn2_6/prodata/2/paa/US40_ COMB.pcp.*	Sequence 19, Appli
37:	/cgn2_6/prodata/2/paa/US41_ COMB.pcp.*	Sequence 20, Appli
38:	/cgn2_6/prodata/2/paa/US42_ COMB.pcp.*	Sequence 21, Appli
39:	/cgn2_6/prodata/2/paa/US43_ COMB.pcp.*	Sequence 21, Appli
40:	/cgn2_6/prodata/2/paa/US44_ COMB.pcp.*	Sequence 23, Appli
41:	/cgn2_6/prodata/2/paa/US45_ COMB.pcp.*	Sequence 2, Appli
42:	/cgn2_6/prodata/2/paa/US46_ COMB.pcp.*	Sequence 2, Appli
43:	/cgn2_6/prodata/2/paa/US47_ COMB.pcp.*	Sequence 3, Appli
44:	/cgn2_6/prodata/2/paa/US48_ COMB.pcp.*	Sequence 4, Appli
45:	/cgn2_6/prodata/2/paa/US49_ COMB.pcp.*	Sequence 5, Appli

# ALIGNMENTS

RESULT 1  
PCT-US03-35733-8  
; Sequence 8, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-8

Query Match 100.0%; Score 3773; DB 1; Length 735;  
Best Local Similarity 100.0%; Pred. No. 5e-299;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	ID	Description
-----	-----	-----	-----	-----	-----

QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDYTFATSDADNHTVMVDDQEVINKASNSKIRLEKGRLYQIKIQY 120  
Db 61 QSAIWSGFIKVKKSDYTFATSDADNHTVMVDDQEVINKASNSKIRLEKGRLYQIKIQY 120  
QY 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSPEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETTKPDMLKEALKIAGFNEPNGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETTKPDMLKEALKIAGFNEPNGNL 540  
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNAQNNILIRDKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNAQNNILIRDKRFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIPSKKGYEIG 735  
Db 721 IKKILIPSKKGYEIG 735

RESULT 2  
US-09-848-909A-8  
; Sequence 8, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; PRIOR FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis

US-09-848-909A-8  
Query Match 100.0%; Score 3773; DB 23; Length 735;  
Best Local Similarity 100.0%; Pred. No. 5e-299;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKKSDYTFATSDADNHTVMVDDQEVINKASNSKIRLEKGRLYQIKIQY 120  
Db 61 QSAIWSGFIKVKKSDYTFATSDADNHTVMVDDQEVINKASNSKIRLEKGRLYQIKIQY 120  
QY 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYMTDSQNKKEVSSDNLOLPELKQKSSNSRKRSTASGTPVDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSPEKVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKSTASDPYSPEKVT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVGNQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVYGNATYNFENGRVVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETTKPDMLKEALKIAGFNEPNGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERBRIAAVNPSPDLETTKPDMLKEALKIAGFNEPNGNL 540  
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNAQNNILIRDKRFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNAQNNILIRDKRFHYDR 600  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNISSLRODGKTFIDFKYNDKPLIYISNPNYKVVAVTKENTIIINPSENGDTSTNG 720  
QY 721 IKKILIPSKKGYEIG 735  
Db 721 IKKILIPSKKGYEIG 735

RESULT 3  
PCT-US03-35733-7  
; Sequence 7, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-7

Query Match      99.9%; Score 3769; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.1e-298;
Matches 734; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPVPDRDN 180
QY 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240
Db 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTPTMNYNQFLELEKTKQLRLDPTDQVYGNATYFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTPTMNYNQFLELEKTKQLRLDPTDQVYGNATYFNGRVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQKDIITEFDFNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDFNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 4
US-09-848-909A-7
; Sequence 7, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A

; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-7

Query Match      99.9%; Score 3769; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.1e-298;
Matches 734; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDNLFQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPELKQKSSNSRKRSTAGPVPDRDN 180
QY 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240
Db 181 DGIPODSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTSKNTSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTPTMNYNQFLELEKTKQLRLDPTDQVYGNATYFNGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTPTMNYNQFLELEKTKQLRLDPTDQVYGNATYFNGRVRVDTGNSWSEV 480
QY 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQKDIITEFDFNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
Db 541 QYQKDIITEFDFNFQOQTSQNIKNQLAELNATNIYTVLDKIKLNKNNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIIEIDTEGLKEVINDRY 660
QY 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 5
PCT-US03-35733-6
; Sequence 6, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
```

; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-6

Query Match 99.9%; Score 3768; DB 1; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.3e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

Qy 61 QSAIWSGFIKVKSDSEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

Qy 121 QRENTEKGLDFKLYWTDSONKEVISSDNLQLPKOKSSNSRKRSTAGTPVPRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKEVISSDNLQLPKOKSSNSRKRSTAGTPVPRDN 180

Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240  
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300

Qy 301 SEVHGNAEVHAGFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAGFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480

Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540  
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540

Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFFHYDR 600

Qy 601 NNIAVGADESIVKAEHREVINSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADESIVKAEHREVINSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660

Qy 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDSTNG 720  
Db 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDSTNG 720

Qy 721 IKKILIFSKKGYBIG 735  
Db 721 IKKILIFSKKGYBIG 735

RESULT 6

US-09-848-909A-6  
; Sequence 6, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-6

Query Match 99.9%; Score 3768; DB 23; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.3e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPSENQYF 60

Qy 61 QSAIWSGFIKVKSDSEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDSEYFATSADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

Qy 121 QRENTEKGLDFKLYWTDSONKEVISSDNLQLPKOKSSNSRKRSTAGTPVPRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKEVISSDNLQLPKOKSSNSRKRSTAGTPVPRDN 180

Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240  
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKWSASDPSDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNTSRTHT 300

Qy 301 SEVHGNAEVHAGFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHAGFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYVNLPTTSLVKGKQTLATIKAKENQLSQILAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNENGRVVRVDTGSNWSEV 480

Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540  
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPKNL 540

Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYVLDKIKLNKAMNILLRDKRFFHYDR 600

Qy 601 NNIAVGADESIVKAEHREVINSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADESIVKAEHREVINSTEGLLNIDKDKIRKILSGYIVIEDETEGLKEVINDRY 660

Qy 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDSTNG 720  
Db 661 DMLNISLRQDGKTFIDFKKYNDKPLIYISNPNYKVVAVTKENTIINPSENGDSTNG 720

Qy 721 IKKILIFSKKGYBIG 735

Db 721 IKKILFSGKGYEIG 735  
RESULT 7  
PCT-US03-35733-5  
; Sequence 5, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; PRIOR FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-5  
Query Match 99.8%; Score 3767; DB 1; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.6e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
Qy 121 QRENTEKGLDFKLYWTDSONKEVITSSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKEVITSSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180  
Qy 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKMWSTASDPYDFEKT 240  
Db 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKMWSTASDPYDFEKT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSTRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSTRTISKNSTSRHT 300  
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGNSWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGNSWSEV 480  
Qy 481 LPOIQETTARIIFNGKDLNIVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPOIQETTARIIFNGKDLNIVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Qy 661 DMLNISSLRQDGKTFDFKYNKDKPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720

Db 661 DMLNISSLRQDGKTFDFKYNKDKPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720  
Qy 721 IKKILFSGKGYEIG 735  
Db 721 IKKILFSGKGYEIG 735  
RESULT 8  
PCT-US03-35733-10  
; Sequence 10, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR APPLICATION NUMBER: US 60/424,987  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-10  
Query Match 99.8%; Score 3767; DB 1; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.6e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSADNHVTWVDDQEVINKASNSKIRLEKRLYQIKIY 120  
Qy 121 QRENTEKGLDFKLYWTDSONKEVITSSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKEVITSSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180  
Qy 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKMWSTASDPYDFEKT 240  
Db 181 DGIPOSLVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKMWSTASDPYDFEKT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSTRTISKNSTSRHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSTRTISKNSTSRHT 300  
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGNSWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNFNGRVRVDTGNSWSEV 480  
Qy 481 LPOIQETTARIIFNGKDLNIVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
Db 481 LPOIQETTARIIFNGKDLNIVERIAAVNPSPDLETTKPDMTLKEALKIAGFNEPENG 540  
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYVLDKIKLNKAKNILLIRDKRPHYDR 600  
Qy 601 NNIAVGADESVMKEAHEVINSSTEGLLINIDKIRKILSGYIVIEDETEGLKEVINDRY 660

```
Db 601 NNIAVGADESUVKEAREVINSSTGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKENTIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKENTIINPSENGDTSNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 9
PCT-US03-35733-20
; Sequence 20, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-20

Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Qy 61 QSAIWGCFIKVKKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKY 120
Db 61 QSAIWGCFIKVKKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVLPVTSVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVLPVTSVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNNYQFLEBKTQKLRDLDQVYGNATYNFENGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNNYQFLEBKTQKLRDLDQVYGNATYNFENGRVRVDTGNSWSEV 480
Qy 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Db 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Qy 541 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNIILIRKRFHYDR 600
```

```
Db 541 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNIILIRKRFHYDR 600
Qy 601 NNIAVGADESUVKEAREVINSSTGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAREVINSSTGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660
Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKENTIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNYKVNVAVTKENTIINPSENGDTSNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 10
PCT-US03-35733-21
; Sequence 21, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-21

Query Match 99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Db 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60
Qy 61 QSAIWGCFIKVKKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKY 120
Db 61 QSAIWGCFIKVKKSDSYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLOLPELKQKSSNSRKRSTAGTVPDRDN 180
Qy 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRHT 300
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTISKNTSSTRHT 300
Qy 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYVNTGTAPIYVLPVTSVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVLPVTSVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNNYQFLEBKTQKLRDLDQVYGNATYNFENGRVRVDTGNSWSEV 480
Db 421 LNAQKDFSSPTITMNNYQFLEBKTQKLRDLDQVYGNATYNFENGRVRVDTGNSWSEV 480
Qy 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Db 481 LPOIQTETARIIFNGKOLNIVERRIAANVPSDPLETTKPDMTLKEALKAFGNPNGL 540
Qy 541 QYQKDIETEFDFNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAKQNIILIRKRFHYDR 600
```



Db 481 LPOIQTETARIIFNGKDLNLFERRIAAANPSDPLETTKPDMLTKEALKIAFGNEPENG 540  
QY 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAQWILIRDKRPHYDR 600  
Db 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAQWILIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

RESULT 11  
US-09-791-537-43735  
; Sequence 43735, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 43735  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: pdb lacc  
US-09-791-537-43735

Query Match 99.8%; Score 3767; DB 22; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.6e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPLPELKQSSNRKKRSTAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPLPELKQSSNRKKRSTAGTPVDRDN 180  
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKSTASDPYDFEKT 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300  
QY 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVLPITSLVIGKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVLPITSLVIGKQNTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVGNATYVNFENGRVVDGTSNWEV 480

Db 421 LNAQKDFSSPTITWYVNFQLEKTKQLRLDQVGNATYVNFENGRVVDGTSNWEV 480  
QY 481 LPOIQTETARIIFNGKDLNLFERRIAAANPSDPLETTKPDMLTKEALKIAFGNEPENG 540  
Db 481 LPOIQTETARIIFNGKDLNLFERRIAAANPSDPLETTKPDMLTKEALKIAFGNEPENG 540  
QY 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAQWILIRDKRPHYDR 600  
Db 541 QYQKDIITFDNFDOQTSONIKQLAELNATNIYTVLDKIKLNAQWILIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKGYEIG 735  
Db 721 IKKILIFSKGYEIG 735

RESULT 12  
US-09-848-909A-5  
; Sequence 5, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-5

Query Match 99.8%; Score 3767; DB 23; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.6e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGFFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDDEVTFATSDNHNVTMWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPLPELKQSSNRKKRSTAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQPLPELKQSSNRKKRSTAGTPVDRDN 180  
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKSTASDPYDFEKT 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSSPEKSTASDPYDFEKT 240  
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRTH 300  
QY 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFFDIGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Db  
361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLKEALKIAFGNEPKNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLKEALKIAFGNEPKNL 540  
Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
Qy 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735  
RESULT 13  
US-09-848-909A-10  
; Sequence 10, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848, 909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201, 800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-10  
Query Match 99.8%; Score 3767; DB 23; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.6e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60  
Db 1 EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60  
Qy 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDDQEVINKASNKIRLEKRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDDQEVINKASNKIRLEKRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSSNRKRGSTAGTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSSNRKRGSTAGTVPDRN 180  
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTADPYSDPEKVT 240  
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTADPYSDPEKVT 240  
Qy 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDOSTONTSETRTISKTSTSRHT 300  
Db 241 GRIDKNVSPARHPLVAAPYIVHVDMENIILSKNEDOSTONTSETRTISKTSTSRHT 300

301 SEYHGNAEVHASFDDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEYHGNAEVHASFDDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVNTGTAPIYNNVLTPTSLVLGKQTLATIKAKENQSLQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
Db 421 LNAQKDFSSPTITMNNYQFLEKTKQLRLDQVYGNATYNFENGVRVDTGNSWSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLKEALKIAFGNEPKNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIIAAVNPSPLETTKPDMLKEALKIAFGNEPKNL 540  
Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKXNMLIRDKRPHYDR 600  
Qy 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735  
RESULT 14  
US-09-848-909A-20  
; Sequence 20, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Sellman, R. John  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848, 909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201, 800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-20  
Query Match 99.8%; Score 3767; DB 23; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.6e-298;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60  
Db 1 EVKQENRLNESSESSQGLLYGYPFDLNFQAPMVVTSSTTGDLSIPSELENIPSENYF 60  
Qy 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDDQEVINKASNKIRLEKRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDYFTFATSADNHVTMVDDQEVINKASNKIRLEKRLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSSNRKRGSTAGTVPDRN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSSNRKRGSTAGTVPDRN 180  
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTADPYSDPEKVT 240  
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKMWSTADPYSDPEKVT 240

Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERWAEATMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERWAEATMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYFPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYFPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSNWEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSNWEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Qy 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 15
US-09-848-909A-21
; Sequence 21, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002 US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-21

Query Match 99.8%; Score 3767; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.6e-298;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVYTSSTGDLSPSSSELENIPSENQYF 60
Db 1 EVKQENRLNESESSQGLGYFSDLNFPQAPMVYTSSTGDLSPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFTKVKKSDEYTFATADNHVTWVDDQEVINKASNSNKIRLEKRLYQIKIY 120
Db 61 QSAIWSGFTKVKKSDEYTFATADNHVTWVDDQEVINKASNSNKIRLEKRLYQIKIY 120
Qy 121 ORENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKRKSTASGPTVPDRDN 180

Db 121 ORENPTKGLDFKLYWTDSONKKEVISSDNQLPELKQKSSNRKRKSTASGPTVPDRDN 180
Qy 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKVT 240
Db 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERWAEATMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERWAEATMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYFPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKAKENQLSQILAPNNYFPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSNWEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYFNFENGVRVDTGNSNWEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540
Qy 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNILLIRDKRPHYDR 600
Qy 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAAREVINSSTGLLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 720
Db 661 DMLNSSLRODQKTFIDFKKYNKDLPLYISNPNYKVNVAATYKNTIINPSENGDTSTNG 720
Qy 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

Search completed: May 3, 2004, 19:57:37
Job time : 190.546 secs

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 seconds

(without alignments)  
5403.004 Million cell updates/sec

Title: US-09-848-909A-8

Perfect score: 3773

Sequence: 1 EVKQENRLNSESSESSQGLL.....TSTNGIKILIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: piri:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3767	99.8	764	I39934	protective antigen
2	877	23.2	875	I40862	iota toxin compo
3	235.5	6.2	192	I39933	cryptic protein -
4	235.5	6.2	204	G59104	hypothetical prote
5	205	5.4	4688	F82885	hypothetical prote
6	187.5	5.0	2401	T28576	rhostry protein -
7	185.5	4.9	1629	S05503	major merozoite su
8	185	4.3	2259	T28577	rhostry protein -
9	184.5	4.9	2529	B64535	toxin-like outer m
10	183.5	4.9	1125	E90598	membrane nucleas
11	183	4.9	4152	T31102	filamentous hemag
12	182	4.8	6713	B89921	hypothetical prote
13	180	4.8	1193	S68218	botulinum neurotox
14	179	4.7	1227	C97033	uncharacterized pr
15	178.5	4.7	1072	A86827	hypothetical prote
16	178	4.7	1302	JC6009	surface-located me
17	178	4.7	4919	T31105	hypothetical prote
18	177.5	4.7	821	S67087	hypothetical prote
19	177	4.7	752	G90599	hypothetical prote
20	177	4.7	1635	A10452	hemolysin [importe
21	175	4.6	1658	S55101	hypothetical prote
22	174.5	4.6	1385	T30822	impli protein Myc
23	173	4.6	5005	F82884	hypothetical prote
24	172.5	4.6	2399	H71879	toxin-like outer m
25	172	4.6	1939	T18372	repeat organellar
26	171	4.5	2178	S55805	alpha-toxin - Clos
27	170	4.5	1631	1 SAZOK1	major merozoite su
28	170	4.5	1837	T41023	probable nuclear p
29	169.5	4.5	1230	S56850	SMC1 protein homol

30	169.5	4.5	2340	2	B71704	cell surface anti
31	168	4.5	1315	2	T28679	fibrinogen-binding
32	168	4.5	1957	2	T38077	hypothetical coile
33	166.5	4.4	1051	2	T18351	impli protein - Myc
34	166.5	4.4	1802	2	S52611	tyb protein - yeas
35	166.5	4.4	2334	2	S32920	cell wall-associat
36	166	4.4	3724	2	T18427	hypothetical prote
37	165.5	4.4	786	2	T18469	hypothetical prote
38	165.5	4.4	979	2	JQ0894	plis protein - Myc
39	165.5	4.4	1115	2	T41342	probable coiled-co
40	165	4.4	1308	2	E71622	probable membrane
41	165	4.4	3216	2	C90538	hypothetical prote
42	164	4.3	1465	2	S31262	tyb protein - yeas
43	164	4.3	1516	2	E71619	RAD2 endonuclease
44	164	4.3	1803	2	S56894	tyb protein - yeas
45	163.5	4.3	624	2	PC6003	surface membrane p

#### ALIGNMENTS

##### RESULT 1

I39934

Protective antigen precursor - Bacillus anthracis plasmid

C:Species: Bacillus anthracis

C:Date: 19-Jul-1996 #sequence,revision 19-Jul-1996 #text\_change 01-Dec-2000

C:Accession: I39934; S69160; F59104

R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.

Gene 69, 287-300, 1988

A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr

A:Reference number: I39933; MUID:89172073; PMID:3148491

A:Accession: I39934

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-764 <RES>

A:Cross-references: GB:M22589; NID:G143280; PIDN:AAA22637.1; PID:G143282

R:Friedman, T.C.; Gordon, V.M.; Leppla, S.H.; Klimpel, K.R.; Birch, N.P.; Loh, Y.P.

Arch. Biochem. Biophys. 316, 5-13, 1995

A:Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SF

A:Reference number: S69160; MUID:95142670; PMID:7840657

A:Accession: S69160

A:Molecule type: protein

A:Residues: 197-202 <FRI>

R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koshle

J. Bacteriol. 181, 6509-6515, 1999

A:Title: Sequence and organization of pXOI1, the large Bacillus anthracis plasmid harbor

A:Reference number: A59091; MUID:99445483; PMID:10515943

A:Accession: F59104

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-313 'O', 315-764 <OKI>

A:Cross-references: GB:AF065404; NID:G4894216; PIDN:AAD32414.1; PID:G4894326

A:Experimental source: strain Sterne

A>Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid

C:Genetics:

A:Gene: pXOI-110

A:Genome: plasmid

C:Function:

A:Description: three component exotoxin; protective antigen binds to receptors on the s

y active components edema factor or lethal factor; the complex is internalized by recep

C:Keywords: exotoxin

F1-29/Domain: signal

F30-196/Domain: propeptide #status predicted <PRO>

F197-202/Product: protective antigen #status experimental <MAT>

Query Match 99.8%; Score 3767; DB 2; Length 764;

Best Local Similarity 99.9%; Pred. No. 1.7e-182;

Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYSFDLNFQAPMVVVTSTTGDLSIPSELENIPSENQYF 60

Db 30 EVKQENRLNSESSESSQGLLYGYSFDLNFQAPMVVVTSTTGDLSIPSELENIPSENQYF 89

```

QY 61 QSAIWSGFTKVKKSDEYTFATSDADNHVTMMVDDQEVINKASNSNKIRLEKRLYQIKIQ 120
DB 90 QSAIWSGFTKVKKSDEYTFATSDADNHVTMMVDDQEVINKASNSNKIRLEKRLYQIKIQ 149
QY 121 QRENPTKGLDFKLYWTDSONKKEVYSSDNLQPELKQKSSNSRKRSTAGTPVDRDN 180
DB 150 QRENPTKGLDFKLYWTDSONKKEVYSSDNLQPELKQKSSNSRKRSTAGTPVDRDN 209
QY 181 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYDFEKT 240
DB 210 DGIPOSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYDFEKT 269
QY 241 GRIDKNVSEARHPLVAAPYVHVDMENIILSKNEDQSTONTSTRIISKNSTSRHT 300
DB 270 GRIDKNVSEARHPLVAAPYVHVDMENIILSKNEDQSTONTSTRIISKNSTSRHT 329
QY 301 SEVHGNAEVAHAFDFTGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 330 SEVHGNAEVAHAFDFTGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 389
QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQIILAPNNYPSKNLAPIA 420
DB 390 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQIILAPNNYPSKNLAPIA 449
QY 421 LNAQKDFSSPTIMYNNQFLEKTKQLRLDQVVGNIATYNFENGVRVDTGNSWSEV 480
DB 450 LNAQKDFSSPTIMYNNQFLEKTKQLRLDQVVGNIATYNFENGVRVDTGNSWSEV 509
QY 481 LPOIQTETARIIFNGKOLNIVERIAAVNPSPLETTKPDMLTKEALKIATFNGPENGML 540
DB 510 LPOIQTETARIIFNGKOLNIVERIAAVNPSPLETTKPDMLTKEALKIATFNGPENGML 569
QY 541 QYQKDIETDFDFNDQOTSONIKNQLAELNATNIYTVLDKIKLNAQNNILIRDKRPHYDR 600
DB 570 QYQKDIETDFDFNDQOTSONIKNQLAELNATNIYTVLDKIKLNAQNNILIRDKRPHYDR 629
QY 601 NNTAVGADESVAKEAREVINSSTEGLLNIDKIDIRKILSGYIVETEDTEGLKEVINDRY 660
DB 630 NNTAVGADESVAKEAREVINSSTEGLLNIDKIDIRKILSGYIVETEDTEGLKEVINDRY 689
QY 661 DMLNLSLRDQGTFFDFKYNKDLPLIYISNPNKYVNVYAVTKNTIINPSENGDTSTNG 720
DB 690 DMLNLSLRDQGTFFDFKYNKDLPLIYISNPNKYVNVYAVTKNTIINPSENGDTSTNG 749
QY 721 IKKILIFSKGYEIG 735
DB 750 IKKILIFSKGYEIG 764

RESULT 2
I40862
iota toxin component Ib - Clostridium perfringens
C/Species: Clostridium perfringens
C/Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Oct-1999
C/Accession: I40862; S42774
R/Perelle, S.; Gilbert, M.; Boquet, P.; Popoff, M.R.
Infect. Immun. 61, 5147-5156, 1993
A/Title: Characterization of Clostridium perfringens iota-toxin genes and expression in
A/Reference number: I40861; MUID:94041637; PMID:8225592
A/Accession: I40862
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-975 <RES>
A/Cross-references: EMBL:X73562; NID:929031; PIDN:CAA51960.1; PID:g414655

Query Match 23.2%; Score 877; DB 2; Length 875;
Best Local Similarity 31.2%; Pred. No. 1.1e-36;
Matches 253; Conservative 131; Mismatches 277; Indels 150; Gaps 31;

QY 1 EVKQENRLNLSSESSQGLGYTFSDLNQAPMVVTVSTTGDLSIPSELENIPSE-NQY 59
DB 35 DTNQKEETNTNLTSSNGLMGYTFADBEFKDLLELWAPIKNGDLKFEKKYDKLLTENDSS 94

```

```

QY 60 FQSAIWSGFTKVKKSDEYTFATSDADNHVTMMVDDQEVINKASNSNKIRLEKRLYQIKIQ 119
DB 95 IKSIRWTGRIIPSEDEGEYILSDR-NDVLMQINAKGDIK-...TLKYNMKGQAYNIRIE 150
QY 120 QREN-----PTEKGLDFKLYWTDSONKKEVYSSDNLQPELKQKSSNSRKRST 170
DB 151 IODKNLGSIDNLVSP-----KLYW-ELNGNKTVIPEENLFFRDYSKIDEND----- 195
QY 171 AGTVP-----DRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIH 211
DB 196 --PFIPNNNFVDRFFSAAWEDBDLTDNDNIPDAYEKNGYII-----KDSIAVKNWDSFA 249
QY 212 EKEGLTKYKSSPEKSWSTASDPYDFEKTGRIDKNVSEARHPLVAAPYVHVDMENIIL 271
DB 250 E-QYKYYSSYLESTAGDPYTDYQKASGIDKAIKLEARDPLVAAPYVGVGMENLII 308
QY 272 SKNEDQSTQNTDSETRITISKNSTSTSTHSEVHGNAEVAHAFDFTGGSVSAGFSNSNST 331
DB 309 STNEHASS-----DQKTVSRATNTSKTDANTV-----GVSTISAGVQNGFTGN 351
QY 332 VAIDHS-----LSLAGERTWAETMGLNTADTARLNANIRVNTGTAPIYVNLPTTSL 383
DB 352 ITTSYSHTTDNTSIAVQDSNGESWNTGLSINKGESAVINANVRVYNTGTAPMYKVTPTNL 411
QY 384 VLGNQOTLATIKAKENQLSQIILAPNNYPSKNLAPIALNAQKDFSSPTIMYNNQFLE 443
DB 412 VL-DGETLATIKADQNGQIGNNLSFNETYPKKGLSPLALNTMDQFNARLIPINDYDQKLLD 470
QY 444 KTKQLRLDQVVGNIATYNFENGVRVDTGNSWSEVLPQIQTETARIIFNGKOLNIVER 503
DB 471 SGKQIKLETTQVSGNYCTKN-SQOI-ITGNSKSNVISQIDSVSASIILD-TGSQTFER 527
QY 504 RIAAVNSDPLETTKPDMLTKEALKIATFNGPENGMLQY-OGKDIET--FDNFDOOTSQ 560
DB 528 RVAAKEQGNPEDKT-PEITIGEAIKKAFSATK-NGELLYFNGPIDESCVELIFDDNTSE 585
QY 561 NIKNQLAELNATNIYTVLDKIKLNAQNNILIRDKR--HYDR-NNTAVGADESVAKEARH 617
DB 586 IIEQKLYLDDKLIYV---KLERGNVILIKVPSYFTNEDEYNFP--ASWSNIDTGNQ 639
QY 618 EVINSSTEGL-----LLNIDKIDIRKILSGY-----IVIEIETEGLEKVIN 658
DB 640 DGLQSVANKLSGETKIIIPMSKLPKRYVFGSGYKSDPSTNSITVNIKSKEQKTDYLV 699
QY 659 RYDMLNLS-----SLRDQGTFFDFKYNKDLPLIYISNPNKYV----- 696
DB 700 EKDYTKSYEFETGKDSIDIEITLTSSGVIFLDNLSITELNSTPEILKEPEIKVPSDQE 759
QY 697 -----NVYAVTKNTIINPSENGDTSTNGI 721
DB 760 ILDAHNNKYADIKLDT-----NTGNTVIDGI 785

RESULT 3
I39933
cryptic protein - Bacillus anthracis
C/Species: Bacillus anthracis
C/Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 15-Oct-1999
C/Accession: I39933
R/Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.
Gene 69, 287-300, 1988
A/Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr
A/Reference number: I39933; MUID:89172073; PMID:3148491
A/Accession: I39933
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-192 <RES>
A/Cross-references: GB:M22589; NID:g143280; PIDN:AAA22636.1; PID:g143281

Query Match 6.2%; Score 235.5; DB 2; Length 192;
Best Local Similarity 34.6%; Pred. No. 2.6e-05;
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

```

QY 587 MNILIRKDFHYDRNNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEI 646  
Db 1 MNILVRDP-YHYDNGNIVGVDDSYLKAYKQILNWSGVSLENLDEVDVQALSGYMLQI 59  
QY 647 EDTE-----GLKEVINDRYDMNMISSLRQDKTFIDPKKYNKDLPLYSNPN 693  
Db 60 KKPSNHLNTPVTITLAGKDSGVGLYRVL-----DGAFLDFNKFENWRSILV-DPG 112  
QY 694 YKVNYYAVTKEN-TIINPSENGDTSTNGIKKILIFSKKGYEI 734  
Db 113 DDVYVYAVTKEDFNATVDRDENGNI-NKLKNTLVLSGKIKEI 153  
RESULT 4  
G59104  
hypothetical protein pX01-111 - Bacillus anthracis virulence plasmid pX01  
C:Species: Bacillus anthracis  
C>Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 11-May-2000  
C:Accession: G59104  
R:Okimaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler  
J. Bacteriol. 181, 6509-6515, 1999  
A:Title: Sequence and organization of pX01, the large Bacillus anthracis plasmid harbored  
A:Reference number: A59091; MUID:99445483; PMID:10515943  
A:Accession: G59104  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-204 <OKI>  
A:Cross-references: GB:AF065404; NID:94894216; PIDN:AA032415.1; PID:94894327  
A:Experimental source: strain Sterne  
A>Note: similar to hypothetical protein in the protective antigen domain; ypa, plasmid p  
C:Genetics:  
A:Gene: pX01-111  
A:Genome: plasmid  
Query Match 6.2%; Score 235.5; DB 2; Length 204;  
Best Local Similarity 34.6%; Pred. No. 2.9e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;  
QY 587 MNILIRKDFHYDRNNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEI 646  
Db 1 MNILVRDP-YHYDNGNIVGVDDSYLKAYKQILNWSGVSLENLDEVDVQALSGYMLQI 59  
QY 647 EDTE-----GLKEVINDRYDMNMISSLRQDKTFIDPKKYNKDLPLYSNPN 693  
Db 60 KKPSNHLNTPVTITLAGKDSGVGLYRVL-----DGTGFLDFNKFENWRSILV-DPG 112  
QY 694 YKVNYYAVTKEN-TIINPSENGDTSTNGIKKILIFSKKGYEI 734  
Db 113 DDVYVYAVTKEDFNATVDRDENGNI-NKLKNTLVLSGKIKEI 153  
RESULT 5  
F82885  
hypothetical protein U0482 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: F82885  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
A:Reference number: A82870  
A:Accession: F82885  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-4688 <GLA>  
A:Cross-references: GB:AE002145; GB:AF222894; NID:96899476; PIDN:AAF30894.1; GSPDB:GN001  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: U0482  
A:Genetic code: SGC3  
Query Match 5.4%; Score 205; DB 2; Length 4688;  
Best Local Similarity 21.3%; Pred. No. 0.1;

Matches 174; Conservative 132; Mismatches 324; Indels 186; Gaps 40;  
QY 19 LLGYFSDNFOAPMVVTSSTTGLSIPSELENIPSENQYFQSAIWSGFIKVKKSDYET 78  
Db 3699 LVDVYLD-NIHQNDITETKIFKQHN-SKEIINPGVTWISKHGWKSPDTDTTANFEK 3756  
QY 79 FATSADNHTVMVDDDEVINKASNKIRLEKELYQIKIQYQRENPTKGLDFKLYWTD 138  
Db 3757 IETQ-----DDNVLNNIDATVKFDEHNNTKQKIVRIKEN-----ND 3795  
QY 139 SQNKKEVISSDNLQLPKOKSN-----SRKKRSTAGPTVDPDR----- 179  
Db 3796 WLKQGI---DNLN-PETKYKLENIELSKPLKTHNTLSVSINDKENISLITETGNPVLKV 3851  
QY 180 ----NGIPDSLEVEGYTDVVKRFTLSPWISNIHEKKGLTKYKSPKSWSTASDPYSD 235  
Db 3852 IQQNDTINDTQQTINVTLSGVNSK-YNGROIKVYKDNNNVIYESS---LITLQKGDND 3907  
QY 236 FEKVTGRIDKNSPEARHPLVAAPVHVDMENILSKN-EDQSTQNTDSET-RTISKNT 293  
Db 3908 YQLLENLSN-----REYFEKIEIHNISNTNFEDEKLGVSNTFITQNT 3957  
QY 294 STSRTHS-EVHGNAEVAHF-----FDIGGSVSAGFS-----SNS---STVAID 335  
Db 3958 TVQWNDSSATVGTGRGVNFNKIKSBDKILENNQVVAWFAPKETIRDTNTLQYTRPLK 4017  
QY 336 HSLSLAGERTWAEWTGLNT---ADTAELNANIRYVNTGTAPIVNLPTTSLVLGKQNTLA 392  
Db 4018 DVTSDFKEGWADLNSVNFKEETTYKLVKIQFVNPKAKNNINNSVILDTNNSI 4077  
QY 393 -----TIKAKENQLSQILAPNNYPSKNLAPIALN-AQDFSSPTITWNY--NQFLELE 443  
Db 4078 NSNYEFTTKVDGDKHLNITSSNNVNTNSQTNFTLSGVKKSVMGKKIKLSYKSNDSSES 4137  
QY 444 KTQQLRLDTPQVYGNATYVNFNGR-----VEVDGSGNSWSEVLPOIQETTARIIFNGKD 497  
Db 4138 HTNEVLIESNKTQVNLNLLNKKRNYTLIDVLDKIDNNVSDPFKSGNLNFSITRTSA 4197  
QY 498 LNLVERRIAAVNSDPLETT-----KPDMLKEA-----LKIAPGFNEPGLNQ 541  
Db 4198 INVNLIEISNRASNTLKSIIKINLNDPDVLRDQDQATIVYGNKNQAMGFITVSGNIK 4257  
QY 542 YQKQDITEFPNPDQTSQNIK-NQLAELNATNIYVLDKIKLNKNNILIRDKRPHYD- 599  
Db 4258 YLTATLYDLNFN-DKVNIVNISFNKPFSAEAEN-----IGDKSNNII-----YNNDS 4304  
QY 600 -----RNNTAVGA---DESVWKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTE 650  
Db 4305 IPKLEINNDIIVNGPINKKEIVVKNANQX--NNIDVDLGLQINPKIAHLN-FAKPKSTN 4361  
QY 651 GLKEVINDD--RYDMLNITSSL-ROGKTFIDF-----KKY----- 681  
Db 4362 -----NDIETNVINGSSLVNNDGKTSIRFTLNLKANKLYSLVDVYVYVNNNSNTIVE 4415  
QY 682 NDKLPLYSNPNYKVNYYAVTKENTINPSENGDTS 717  
Db 4416 SNKLP-KLNNINVOIK---INKSHITII--SKNGEWS 4445  
RESULT 6  
T28676  
thoptry protein - plasmodium yoelii (fragment)  
C:Species: plasmodium yoelii  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T28676; A45521  
R:Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.  
Mol. Biochem. Parasitol. 76, 329-332, 1996  
A:Title: Comparison of two members of a multigene family coding for high-molecular mass  
A:Reference number: Z20507; MUID:97077455; PMID:8920022  
A:Accession: T28676  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2401 <SIN>

A;Cross-references: EMBL:U36927; NID:gl041784; PID:gl041785; PIDN:AAB41263.1  
R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-246, 1990  
A;Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co  
A;Reference number: A45521; MUID:91101660; PMID:2270106  
A;Accession: A45521  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 2260-2401 <KE>  
A;Cross-references: GB:M34281

Query Match 5.0%; Score 187.5; DB 2; Length 2401;  
Best Local Similarity 21.7%; Pred. No. 0.29;  
Matches 182; Conservative 118; Mismatches 317; Indels 221; Gaps 43;

QY 1 EVKQENRLNESESSQGLLYYFSDLN-----FOAPMVVSSFTGDLISL 45  
DB 402 EVAKENVQLNVYKSNILEIKXHYNDQINDNTEKEAKQNYDQFKEHMKTIIPPNEWKYQK 461

QY 46 PSELENIPSENYFOSAI--WSGFIKV--KKSDEYTFATSDADNHVTWVDDQEV--I 97  
DB 462 PSIEIKIMKDE--FLSKYKNYNDPKVYKEVSEHNKFTLTWNKITEVSDRIKKYE 518

QY 98 NKASNNKIRLEKGRUYQIKIQORENPTKGLDFKLYWTDSONKKEVISSNLOLPELK 157  
DB 519 NKFNDSKSLINETKK--SIEEYQININTLKKVDD--YIKVCLNTNELITNCHNQOTLK 573

QY 158 QKSNRKK--KRSTAGPTVPDRNDGIDP-----SLEVEGYTVDVKNKFTLSP 205  
DB 574 DKLNQIKITKETSNDIKYTDKFNILDTDKKTELETKTGLSLNHNBSNKKELTYFD 633

QY 206 WISNIEHKKGLTYKSPKASTADPYSDFKVQGRIDKNVSPPEARHPLVAAYPIVHVD 265  
DB 634 LKANLGNKENMLYKOFNEK-----ERAVEDIKKNVDINKIVSN-----IEIT 677

QY 266 MENILSKNEDQSTQNTDSTRITISKNSTSTRTSEVHGNVNAEVSFDDIGGSVAGFS 325  
DB 678 IYTSIYNED-----TENB--IGKSIENLTKVLE-----KVKANVT 713

QY 326 NSNSTVAI--DHSLSLAGERTWAGTGLNADTARLNANIRYNTGTAPIYVLPV 384  
DB 714 NLNEIKELKXDYDFQPGK-----EKNIKYPDEN--KIKNDIDITLQK 754

QY 385 LGKN--QTLATIK--KENQLSQI-----LAPN-----NVYP--SKNLAIALNAQK-- 425  
DB 755 IDKSIEFTLTIKKNSENHIDEIKGQIDKLVKPNKTFMNEDEPKIEKKIENIEVEKIDKKK 814

QY 426 -----DFSSTPIIMYNOFLEKTKQLRLDTPQVGNIAIYNFENGVRVDTGNSWSE 479  
DB 815 NIYKEIDKLLNEISKIENDKTSLEKLNINLSYKSLGNLFLQOIDEKKKAETHIKAME 874

QY 480 V-----LPQIQTETARIIFNGKDLNL-----VERIAAVNPSPD-----LETT-- 518  
DB 975 AYIDDLNLIKKSQEI-----EKEMNINMDIKQDILKEMKALNTSHDDYKITYHTTSKNHEK 931

QY 519 PMDTLKEALKIAFGFNEP-----NGNLOYQGKDITEFDNFQDQTSQNIKNQLAELNATNY 575  
DB 932 ISDIRKNSLXIQDFSESYINDIKKELEKNVLE-----SQNNNTDINQVLSKIE--NIY 984

QY 576 TVLDKIKLNKAMNILLRDKFP--HYDRNNIYAGADESVKAEHREVINSTEGLLLNIDK 633  
DB 985 NIL-----KLNKIKKIIDKREYTDIEKKN-----KKINAELSNS----- 1020

QY 634 DIRKILSGYVEITDEGLKE-----VINDRY-----DMLNISLRQDGKTFIDFKKYN 683  
DB 1021 --EKI-----ITQLKENSLEKOSKIKSTIDNNVSECIKNTNL-----KTIYVNEKN-- 1069

QY 684 KLPAYISN--PNYKVN-----YAVTKNTIINPSENGDSTN--GKKILIFSKK 730  
DB 1070 -INTYFKNABEYNQNSLNFNNIEMADTKSQYILNKQNGTNTNTDYNIKELKEHKK 1126

RESULT 7

S05603  
major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (s  
N;Alternate names: gp195 surface antigen  
C;Species: Plasmodium falciparum  
C;Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 09-Jun-2000  
A;Reference number: A45521; MUID:91101660; PMID:2270106  
A;Accession: S05603; S04850  
R;Myler, P.J.  
submitted to the EMBL Data Library, April 1989  
A;Reference number: S05603  
A;Accession: S05603  
A;Molecule type: mRNA  
A;Residues: 1-1639 <MYL>  
A;Cross-references: EMBL:X15063; NID:g9896; PIDN:CAA33163.1; PID:g9897  
R;Myler, P.J.  
Nucleic Acids Res. 17, 5401, 1989  
A;Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas  
A;Reference number: S04850; MUID:99345116; PMID:2668887  
A;Accession: S04850  
A;Molecule type: mRNA  
A;Residues: 1504-1639 <MYL2>  
A;Cross-references: EMBL:X15063  
C;Superfamily: major merozoite surface antigen  
C;Keywords: glycoprotein; merozoite; surface antigen  
F;1-19/DNA: signal sequence #status predicted <SIG>  
F;20-1639/Product: major merozoite surface antigen #status predicted <MAT>

Query Match 4.9%; Score 185.5; DB 2; Length 1639;  
Best Local Similarity 20.5%; Pred. No. 0.21;  
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

QY 5 ENRL-LNESESSQGLLYYFSDLNFCAPMVVTSITGDLSPSELENIPSENYQFQSA 63  
DB 927 ENILSLGKNKNTYQELIGKQKSE-NF-----YEKILKDSDFTYNE 965

QY 64 IWSGFIKVKSDYTFATSDADNHVTWVDDQEVINKASN-----NKIRLEKGRLY- 114  
DB 966 SFTNFVSKADD-----INSLNDESKRKKLEEDINKLTKTQLSFDLYNKKYKLERLFD 1020

QY 115 -----QIKIQORENPTKGLDFKLYWTDSONKKEVISDNLQLPKOKSNSSKRS 168  
DB 1021 KKKTVGKYKMQIKKLTLLKEQLESKL--NSLNNPKHVL--QNPVSFNFKKGAEIAETEN 1076

QY 169 TSAGPTVPDRNDGIDPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKY--KSSP-- 223  
DB 1077 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1103

QY 224 ---EKWSTADPYSDPE-----KVTRIDKNVSPPEAR-----HPLVAAYPIVHVD 265  
DB 1104 TLSEESIQTEDNYASLENFKVLKLEKGLKNLNLEKKLSYLSGLHLHLLA-----E 1156

QY 266 MENILSKNEDQSTQNTDSTRITISKNSTSTRTSEVHGNVNAEVSFDDIGGSVAGFS 325  
DB 1157 LKEVKNKN--YTGNSPSENNT-----DVNNALSYKKFLPEGTDAIVVS 1200

QY 326 NSNSTVAIDHSLSLAGERTWAGTGLNADTARLNANIRYNTGTAPIYVLPVPTT---- 381  
DB 1201 ESGSDTLTLEQSQPKKPPASTHVGAE--NTITTSQ--NVDDDEVDDVIIIVIFGESEEDYDDL 1256

QY 382 -SLVLGKNQTLATIKAKENQLSQILAPNYYPSKNLAPIA--LNAOKDFSSPTITWYN 437  
DB 1257 GQVTVGAIVPSVI--DNILSKI--ENEYEVLYKPLAGVYRSKQLNENNVMTFVN 1310

QY 438 -----QFLELEKTKQLRLDTPQV--YCNIAI-----YNFENGVRVDTGNSWSEVL 481  
DB 1311 VKDILNSRFNKNFKNV--LESDLIPYKDLTSSNYVVKDPYKFLNKEKROKFLSSYNIK 1369

QY 482 PQIQTETARIIFNGKDLNLVERRIAANVPSDPL-----ETTKPDM--TLKEALKIAGF 533  
DB 1370 DSID-----TDINFA-----NDVLGYKILSEKYSKSDLSIKKINDKQGE 1410

QY 534 NEPN-----GNLQYQKDIIE-----PDFNFDQDQTSQNIKNQLAELNATNY 576  
DB 1411 NEKVLPLFLANNIETLYKTVNDKIDLVFIHLEAKVLNITYEK--SNVEVKIKELN--YLKT 1465

QY 577 VLDKIKLAKQN--ILIRDKRFHYDRNN-----Iavg-ADESVYKAEHREVINSFEGLL 628  
Db 1466 IQKLAADFQKQNNFVGADLSFDYNNHLLTKFLSTGMVFENLAKTVLSNLLDGNLQG-M 1524  
QY 629 LNIIDKD--LRKILSGYIVIEDTEGLKEVINDRYDMLNLSLRQDGKTTIDSKYNDKLP 686  
Db 1525 LNISQHCQVCK-----QCPQSGGCFRLHDE-----RECKCLNTKQBGDKC- 1566  
QY 687 LYSINPNYKYNVAVTKENTIINPSNG-----DTSTNGIKKI 724  
Db 1567 --VENENPTCN-----EN-----NGGCDADAKTEEDSGSNG-KKI 1599

RESULT 8  
T28677  
rhoptry protein - Plasmodium yoelii  
C:Species: Plasmodium yoelii  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C:Accession: T28677; C45521  
R:Keen, J.; Sinha, K.; Brown, K.; Holder, A.  
Mol. Biochem. Parasitol. 65, 171-177, 1994  
A:Title: A gene coding for a high molecular mass rhoptry protein of Plasmodium yoelii.  
A:Reference number: Z20508; MUID:95021522; PMID:7935623  
A:Accession: T28677  
A:Status: preliminary; translated from GB/EMBL/DBU  
A:Molecule type: DNA  
A:Residues: 1-2269 <KEE>  
A:Cross-references: EMBL:L27838; NID:g457145; PID:g457146; PIDN:AAA21304.1  
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42, 241-246, 1990  
A:Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co  
A:Reference number: A45521; MUID:91101660; PMID:2270106  
A:Accession: C45521  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 2131-2269 <KE2>  
A:Cross-references: GB:M34283

Query Match 4.9% Score 185; DB 2; Length 2269;  
Best Local Similarity 19.1%; Pred No. 0.36;  
Matches 161; Conservative 140; Mismatches 328; Indels 214; Gaps 37;

QY 1 EVQENKL-----LNESSSQGLGVYFSDLPNQAPMWVTSSTGDLSPSSSELENIPSE 56  
Db 485 QTLENKFTBFSLNNHEANNELI-KYFSDLKANLGINENMLYNQFTEKEKTFNDIKEK 543  
QY 57 NQYFQSAIWSGFKVKKSDSYTATSADNHVTKWVD-----QEVINKASNSKIRLSK 110  
Db 544 NTHNEISKIEKIHAS-IYNTSEETEREIGINISLNTKVPKEKVENVTNLNKIK-BK 601  
QY 111 GRLYQIKIQRENPTKEGLDFKLYWTDSONK-----REVISSDNLQ----- 152  
Db 602 LKHDF-----SPFGKEGNIKYTDKIKINDDIMAVSQIQDQHINGLDDIQKSES 652  
QY 153 -LPFLQKSSNRKRSSTAGTPVRDNDGIPDSLEVEGYTVDNKRTFLSPWISNTH 211  
Db 653 YVSEMQINKLEKVSNT-----EISND-----NVEG-----IKKQQII---VTKID 692  
QY 212 EKKG-----LTKYKSSPEKWSSTASDPYDFEKTG-----RIDQNSPEARHPL 255  
Db 693 KKKNIYEINKLJSEISKIEKNTSLKVDINLSYGQNLGNLFLSQIDEE-KKKAENYI 751  
QY 256 VAAPYVHVDMENILISKNEQSTQNTDSTRISKNTSTSRTHTEVHGNAEVH----A 311  
Db 752 KSEAYID-DLDNIIKKKSQBIETEMDIKMDINKEMEALKISHDDDKKCHDKSKNHKENS 810  
QY 312 SFFDIGGSVAGFSNNSSTVAIDHSISLAGERTWAETWGLNTADTARLANIYRYVTGT 371  
Db 811 DTDYKSKIIQDFER-ESDINDINKLQ-----KNVSESQNHNSDINQCLNEV 857  
QY 372 APIYNNLPTTSLVLGRKQNTLATIKAKENQLSQILAPNNYFYS-----KNLAPIALNAQ 425

```

858 ANTYNIL-----KLNKKIIDIKKVYTSIERKKNKNDLNNSEK 899
426 -----DFSSTPITVMNQFLE-----LEKTKQLRLDITD-----QVYGN 458
900 VIKKIEGDLSLKECRSKINSLDLDKIDECIKNINVLKKNILNEETNITNHPKNAEYKN 959
459 IATYNEGRVRVDTSNWSEVL PQIQTETARIIFNGKDLNLVERIAAVNPSPDLETTK 518
960 IVLSNFNN-----IEMADNKSQYILEIKKNGT---NDHDYNIKELK-SHKOKSNGYK-TE 1010
519 PDMTLKALKIAFGFNEPKNLQYQGD-----ITEFDNFDDQOTSNIKKNQLAE 568
1011 ADQNKKAIOK-----NKLEFQYKEEVTVLLNKYYAVELKKNKFD-KTKDNKSQIIKE 1061
569 LNATNTYTVLDKTKLNKKNN-----ILIRKRFHYDRNRNIAVGADSVVKEAHRVIN- 621
1062 IKDAHNYCTLESKSEKKNWEIKNEKIHIEDEVANDKSNKAITSIKVSPEPKTKIIKI 1121
622 ---SSTEGLLNIDKDIRKILSGVIVEIEDT-----EGLKEVINDRYDMLNTSSLR 669
1122 NEIRTSDDCKETNLEKQISNLSIDTQTKUTENGKQLKTLLEL-----LESK 1172
670 QDGKTFIDFKKYNDKLPYISN-----PNKYNVYAVTKENTIIINPSENGDTSTNG 720
1173 KQKNKNIEDQKLEDEVNSKIKNIENTVQHKKNYEGIGI-VEKINIATKNKQIESTKE 1230
721 IKK 723
1231 LIK 1233

RESULT 9
B64635
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
R:Accession: B64635
R:Tomb, J.P.; White, O.; Karlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
Petersen, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenn
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: B64635
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2529 *TOM*
A:Cross-references: GB:AE000602; GB:AE000511; NID:g2314060; PIDN:AAD07969.1; PID:g23140
Query Match 4.9%; Score 184.5; DB 2; Length 2529;
Best Local Similarity 21.0%; Pred. No. 0.44;
Matches 158; Conservative 94; Mismatches 261; Indels 241; Gaps 36;
QY 123 ENTETKGLDFKLYWDSQNK-----KEVISSDNQLPELKQSSNRKKRSTSGAGTVPDR 178
DB 729 QNPPAESVSGVYWL--QNKTYNKGYYID----PNLSGGSGSGNTLSTYTNALP--- 779
QY 179 DNDGIPDLSVEGYTVDVKNKRTFLSP---WI-----SNIHEKKGLTKYK 220
DB 780 ---GRSFVNIQNGTLIIGNNTSVNSGLIWGHGGFVITGTFSAANIY---LNNF 832
QY 221 SSPEKWT-----ASD-----PYSDFEKVGTGRIDKNVSEARHPLVAAPIV 262
DB 833 KTEGVNSDGGGANITFRASDMITMDGLNYNDAETVTKMIQTGAS---QHSYATFDALN 889
QY 263 HVDMENILSKNEDQSTQNTDSTRITSKNTSRTSTSEVHGNAEVHAGFPDGGVSVA 322
DB 890 NISVNSSFS-----DMTWGFESFAKNISFS-----NASF-----S 921
QY 323 GFNSNSSTVA---IDHSLSLAGERTWAETWGLNTADTARLNANIRVYVGTAPIYVL- 378
DB 922 GFTNPGGSSVISANATNSLSFINSR-----LNGGAVYNLQANSLEFN-NTQAVFNLY 973

```



379	-----PTSLVLGKQOTLATIKAKENQLSGLIAPNNYPSKNIAPIAL--NAQKDFS	428
974	SRGTSNFNATTOQLGNFTLSS-----OSLLFNFGDTTLONNANITLGNKSQAAPK	1025
429	SPETIMVYNQFLELEKTKOLRLDTPQVYGNIATYFNGRVRVDTGSNWSEVLPOIQETT	488
1026	NS-LFLDNNSNLSLQSVLNANTSAFNNQASLIYNGS-----QATF	1068
489	ARIIFNGKDLNL-VERRIAAVNPSPLETTKPDMLTKEALKIAFGNPNGLVQGGKDI	547
1069	NSLFFNGGTLNASSKLNASPSNNTT--INLDDSVLSASNTSSLNANINPQASQ	1125
548	TEF-----DENFDQOTSONIKNOLA-----E	568
1126	ADFGNNTIITASFPDSASSLNFNNLTANGALNFNGYTPSLTKALMSVGQFVLGNNGD	1185
569	LNATNIYTVLDKIKLNAMNLIIRKRF-----HYDRNNIAV	605
1186	INLSDI-NIFDNIITKSVTYNIIILNAQGITGIGANGYEKILFYGMKIQNIATYSDNNNIQT	1244
606	GA-----DESVVKEAHR-----EVIN--SSTEGLLNIDKDI-----RKILSGYIVE	645
1245	WSFINPLNSQIQESIKNGDLTIEVLNPNPASNTIFNIAPELNYQASKNQPTGYSYD	1304
646	IDETELKEVINDRYDMLNLSI-----BQDGKTFFDKKYNDKPLY---ISNPNY	694
1305	YSDNQA-----GRYLYTSNIKGLFTPKGSQTPQAPGTSPFNQPLSSLIYNYKGFSSENL	1359
695	KUNVYAVTKENTIIIN---PSENGDTSNGIKKIL	725
1360	KTLGLILSONSATLKEMIESNQLDNITN-INEVL	1392

RESULT 10

E90598

membrane nuclease, lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C/Species: Mycoplasma pulmonis

C/Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

C/Accession: E90598

R/Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.; Nucleic Acids Res. 29, 2145-2153, 2001

A/Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis

A/Reference number: A99512; MUID:21267165; PMID:11353084

A/Accession: E90598

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1125 <R>

A/Cross-references: GB:AL445565; PID:g14090108; PIDN:CAC13866.1; GSPDB:GN00153

A/Experimental source: strain UAB CTIP

C/Genetics:

A/Gene: MYPU 6930

A/Genetic code: GSC3

Query Match	4.9%	Score 183.5;	DB 2;	Length 1125;
Best Local Similarity	19.1%	Pred.No. 0.15;		
Matches 165;	Conservative 162;	Mismatches 330;	Indels 207;	Gaps 40;

QY	3	KQENRLNSESSESSQGLGYVFDLNFQAPMWYTSSTTGDLSPSGELENIPESENQVFQS	62
DB	120	ENDSNVNNKENSSSK-----NDEN---LTLTKVLGHNVV-LNQSNIPTKNIATISK	168
QY	63	AIMS-----GFIKY-----KKSEYTFATSAADNHVTMWVDDQE----	97
DB	169	VILHNKLDLVLGLTEITNENGVKTTIINELNKSDDKMYIVSLKKRGTTGSSGQEHVGII	228
QY	98	NKASNSNKIRLE-----KGRLYQIKIQYQRENPTKEGL-----	130
DB	229	YK-----ENKLTLESFDDKEKSGKYEYENKLW---DDPKKGOKIDFVRPPGVKFTKNGI	282
QY	131	--DFKLW--TDSQKKE--VVISDNLQJLPFLKQKSSNGKKRSTKSAGTVPFDRNDGI	183
DB	283	KNDFTVFMHSDAPGVKEERGEISAKGYSGOGHKEVAELRTKEVMEYFDSI-----DGV	337

Db 217 PNGITLVGKVTINTDRFVVST-----DIIPHRENGLLSVNGKVTIDKGGVA 264  
 QY 115 QIKIQY---QRENPEKGLDFKLYWDSQNKKEVISSD-----NLQPELKOKSS 161  
 Db 265 TNGLSHFVAVARNIDQK---KITVAKTENQKSVNPANITFAAGSLNVLNKTREATPISS 321  
 QY 162 NSRKRSTAGPTVDRNDGIPDSLEVEGTVD-----VKKRTFLSPWISNIHEKGLT 217  
 Db 322 QTSR---TSDTPAI--SADSAGSMYGSNIKFVVVTDKAGVGHKGIIFSENDINIKWDGNA 377  
 QY 218 KYKSPKSTASDPYDFEKTGRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQ 277  
 Db 378 SLKELYAK-----KODILAKDIELTEKQLOANNKIILNSTGKINLNASV 425  
 QY 278 STQWDSSTRIS--KNTSSTRTHSEVGNRAEVHASFDFDGSVSAGFSNSNSTVAIDH 336  
 Db 426 SADNVNVSLENALENASMSANSLDVIVTKIEVRS-----SKVSAGTANIKASNIITLD 480  
 QY 337 SLSLAGERTW-----AETWGLNTADTARLN-----ANIRVYV---T 369  
 Db 481 SSVVANKITLVNTNATLNQSKLSAKDMELNVTHNITLNTSKLSAQKANIKTENLTN 540  
 QY 370 GTAPIYNVLPPTSLVLGK-----NOTLA-----TIKAKENQ--LSQILAPNN-- 409  
 Db 541 GEA-----SLVAEKLIDINAIDKITNNGTITAGLTANITTKALENRDNALILAHQNLN 591  
 QY 410 -----YY-----PSKNLAPIALNAOKDFSS-----TPITMYNQFLELEKTKOL 448  
 Db 592 FTVNGSHVYKGDIVSKKAVTFPNSNDSFTSNGSKLVDAQNNLTNVNANNITQGSSEII 651  
 QY 449 RLDTQVYVNI---ATYAFEN--GRVRVDTGNSWSEVLPOIQTETARIIFNGKDLNIVER- 503  
 Db 652 -----LHGNTVLNAKGNFTSNGNLTMMKELNISIESFI---NAGNLTGKLEHVSNT 702  
 QY 504 -----RIAAVPSPLETKPDMTL--KEALKIAGF---FNEPNGNLOVQCKDITEF 550  
 Db 703 TVKNDGLVSIENLNISSKTDFTNNGTLGLEALKIAGGFTNASGSLA--SNKSLDIY 761  
 QY 551 DFNF--DOOTSQNIKNQALNAATNYTVLDK--IKLNKAKMILIRDKRPHYDRNNIAGA 607  
 Db 762 GNNFTNGTIESVKS---LNTNNVTFINNATIKSYGVNLITSQ--NFTNDSNGTVMSH 816  
 QY 608 D-ESVVKAEHREVINSST-----EGLLL-----NIDKDKILSGYIVEIEDTEGLKEVIND 658  
 Db 817 DLNLTISQAN--IINKNLLAGOGLNLTAKGNITNDSNSTAIKVLHNSNDIN-----LNA 869  
 QY 659 RYDMLNLSL--RQDGKTFIDFKKYNKPLPLVYISNPNYKVNVAVTKEINTIINPSENGDTS 717  
 Db 870 NKNYVNIIGIYSQAGNISVEAKLHNDVKL-----SGNITTTTK-----SGNATVK 915  
 QY 718 TNGI 721  
 Db 916 TNSI 919

RESULT 12  
 B89921  
 hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: B89921  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
 A:Reference number: A89758; PMID:21311952; PMID:11418146  
 A:Accession: B89921  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-6713 <KUR>  
 A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BA842527.1; GSPDB:GN00149

A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: ebha  
 Query Match 4.8%; Score 182; DB 2; Length 6713;  
 Best Local Similarity 20.5%; Pred. No. 2.5;  
 Matches 171; Conservative 130; Mismatches 299; Indels 234; Gaps 44;  
 QY 9 LNESESSGGLGYFSDLNFOAPWVITS-----STTGDLSPSSSELENIPSENQYFQSA 63  
 Db 2468 VRQAKSDAKANLG--TLTHLNAQKQDLTSIQEGATTVNGSVNKTQAQDLGAMORLESA 2526  
 QY 64 IWSGFTKVKKSDEY-----TFATSADNHVTWMDVQEVINKASNSKNIRLERGLRYQIKI 118  
 Db 2527 I-ANKDQTKASENYIDADPTKTAFDNAIT---QAESVYLNKDHGTNKDK-----QAVBQ 2576  
 QY 119 QYORENPTKGL--DFKLYWDSQNKKEVISDNQLQPE---LKQKSSNSRKRSTSGAP 173  
 Db 2577 AIQSVTSTENALNGDANLQCAKTEATQAIADNITQLNTPKTALKQOYVNAQR----- 2628  
 QY 174 TVPDRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEK-----KGLTKYKSSPEK 225  
 Db 2629 -----VSGVT--DLKNSATSLNNAQDLKQAIGDHDITVAGGVNVTWASDPK 2672  
 QY 226 WSTASDPYDFEKTGRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSE 285  
 Db 2673 QGAYTDAYNAAKNIIVNG--SPNVITNAADVTAATQRV-----NNAETSLNGDTN 2719  
 QY 286 TRFI---SKNTSSTRTHSEVGNRAEVHASFDFDGSVSA-----GFSNSNSTVAIDHS 337  
 Db 2720 LATAKQAKDALRQMTLSDAQQS-----ITGIDSATQVTGVQSVKDNATNLDNA 2771  
 QY 338 L-----SLAGERTWAEITMGLNTADTARLNANIRYVNTGTAPIYVNLPTLSVLGKNQTL- 391  
 Db 2772 MQLRNSIANKDEKVASQPVVDATDKQNA-----YNTAVTSAENIINATS-----OPTLD 2822  
 QY 392 -ATIKAKENQ-----SQILAPNNYPSKNLAPIA--LNAOKDFSTPITWYNN--Q 438  
 Db 2823 PSATVQAANQVNTNKTALNGAQLNKKOETTANIRLSHLNNAQKQDLNTQVTHAPNIS 2882  
 QY 439 FLELEKTKOLR-----DTQV-----YGNAT-----YNPE 465  
 Db 2883 TVNQVTKAEQLDQAMERLINGIQDKQVQKSVNFTDADPEKQATYNNAVTAENIINQA 2942  
 QY 466 NGRVRVDTGNSWSEVLPOIQTETARIIFNGKDLNIVERRI--AAVNPSPDPLET----- 516  
 Db 2943 NG-----TNANOSQVBAALSTVTTTQALNG-----DRKVTDAKNANQTLSTLDNLNN 2991  
 QY 517 -----TKPDMTLKEALKIAGFNEPNGNLQ--YQKDKITEFDFNFD----- 555  
 Db 2992 AOKGAVTGNINQAHTVAEVT--QAIQTAQELNATMGNLKNLNDKDTLLGSGNQFADADPE 3049  
 QY 556 -----QOTSQNIKNQALNAATNYTVLDKIKLNKAKMILIRDKRPHYDRNNIAGADES 610  
 Db 3050 KKNAYNEAVRNAENILNKSTGTNV-----PKDQVEAAMN-----QVNTTKAALNGTON 3097  
 QY 611 VVKEAHEVINSSTGL--LLNIDKDKIRKILSGYIVEIEDTEGLKEVINDRYDMLNLS- 667  
 Db 3098 L--EKAKQHANTAIIDGLSHLTNAQKALKQLVQOSTTVAEAGNEQKAN-----NVDA 3149  
 QY 668 ---LRQ---DGKTFIDFKKYNKPLPLVYISNPNYKVNVAVTKEINTIINPSEN 713  
 Db 3150 MDKLRQSIADNATTKQNQYTD-----ASPN-KDAYNNAVTTAQGIIDQTTN 3196  
 RESULT 13  
 S68218  
 botulinum neurotoxin type A, nontoxic/nonhemagglutinating - Clostridium botulinum (stra.  
 C:Species: Clostridium botulinum  
 A:Variety: strain NIH  
 C:Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 02-Jun-2003  
 C:Accession: S68218; S74301  
 R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.

FBBS Lett. 376, 41-44, 1995  
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components  
A:Reference number: S67988; MUID:96096783; PMID:8521962  
A:Accession: S68218  
A:Molecule type: DNA  
A:Residues: 1-1193 <FUJ>  
A:Cross-references: EMBL:D67030; NID:G2160224; PIDN:BAAL1050.1; PID:d1011710; PID:g11324  
A:Experimental source: strain NIH  
A:Accession: S74301  
A:Molecule type: protein  
A:Residues: 1-13;145-155 <FUI>  
A:Experimental source: strain NIH  
C:Genetics:  
A:Gene: ant  
C:Superfamily: tetanus toxin  
C:Keywords: neurotoxin  
F:1-1193/Product: botulinum neurotoxin type A, nontoxic/nonhemagglutinating #status expected

Query Match 4.8%; Score 180; DB 2; Length 1193;  
Best Local Similarity 19.6%; Pred. No. 0.25;  
Matches 191; Conservative 131; Mismatches 319; Indels 294; Gaps 46;

QY 42 DLISIP---SSELENTPEQYFQSAIWGFIKVK--KSEYTFATS-----ADNHVT 88  
DB 241 DLIVPYRLRSELENTPEYQALNVLDLLVGGIDPKFINTDPYWFTONYFSNAKKVFEDEHN 300  
QY 89 MWVDDDEVINKASNSNKIRLEK-----GRLYQIKIQY-----QRENPTKGLDF 132  
DB 301 IYETEIEGNAIGNDIKLRKQKFRININDIWEALNLYFSKFFSIMPDPFRFNALKHFYR 360  
QY 133 KLYW-----TDQNKKEVISS-----148  
DB 361 KQYKIDYPENYSINGVNGQINQAULSDRNDQIINKEEIIINLLNGNVSLSMESNYG 420  
QY 149 -----DNLOLP-----ELKQKSNRKRKSTAGTPVPDRDNDGIPDSLEVRGY 192  
DB 421 DGLKSTVDPFYSNKIPYRNAYEYHFNNSDLSLNVNIGVI-----DNIFEIDVNPY 474  
QY 193 TVD-----VKNKRTFLS--PWISNHEKGLTKYKSPKSTASDPVSDFEKVTG 241  
DB 475 KENCDFSPVQKITSTREINTNIPWPNY-----LQANTNNEKFSLS-----SDFEVVS 525  
QY 242 RIDKNSVSPARHPVAAPYVHVHDMENII--LSKNEQDQSTQNTDSB---TRTISKNTST 295  
DB 526 SKDKS-----LVYFSLNVMFYLDISKNSPIDTDKKYILWLEIFRNYSF 571  
QY 296 SRTHTSEVHGNAEVH-----ASFPDGGSVSAGFNSNS--STVAIDHLSLAGEAT 345  
DB 572 DITATQEINTNCGINKVVTWFGKALNLTSDSFVEEFQNLGAISLKNKENLSMPFIIBS 631  
QY 346 WA---ETWGLNTADTARLNANIRYVNTG--TAPIYVNLPT-----TSLVLGKNT 390  
DB 632 YEIPNDMLGFLNDLNEKLENIYSKNTAYPKKIYYNPLDQWTOYYSQYFDLICHAKRSV 691  
QY 391 LA---TIKAKENQLSQILAPNNYPSKNLAPIAL---NAQKPFSS--TPITMN-----435  
DB 692 LAQETLIKRIIOKKLSYLGINSISSDNALMLNTTNTLRDISNESQIAMNVDNFLN 751  
QY 436 -----YNQFLEKTKQLRLDQDQYV--NATYFNGRVRVDTGSNWSEVLQP 483  
DB 752 AAICVFPESNYPKFIISF-----MEQCINNINIKTEF-----IQKTNINEDEKLQ 798  
QY 484 IQETTARIIFNGKDLNLVERRIAAVNPSPLETKPDMTLKEAL-----KIAFGNFPNGN 539  
DB 799 INQN-----VFNSLDFFELN-----IQNKSLFSSSETALLIKEETWPVELVLYAFKEPGNN 849  
QY 540 -----LQVQGDKI--TEFPNFD-----QQTSONIKQLAELNATNITYV-----577  
DB 850 VIGDASGKNTSIEY--SKDIGLVINGSDALYLANGSNQISFSNDFPENGLTNSFSIFYWL 908  
QY 578 --LDKIKLNACKV-----NIIIRD---KRFHY---598  
DB 909 RNLGKDTIKSLKSKEDNCGWEIFQDTGLVFNMDNSNGNEKNIIYLSDSVNSNSWHYITI 968

QY 599 --DR--NNIATGADSESV--KEAHREVINSSTEGLLNIDKIRKILSGVIVIEIDTEGLK 653  
DB 969 SVDRLEKEQLLIFIDNLANVESIKRILNLYSNIISLSENNPSYIEGLTILNKPTTS-Q 1027  
QY 654 EVINDRYDMLNLISSLRQDQKTFIDPKKYNKDLPLVISNPKYKN---VYAVTKENTI---707  
DB 1028 EVLSNFEVLNNSYIRDSNEERLE---YNKTYQLY---NYVFSKPICEVKQNNIYLT 1080  
QY 708 INPSENGDTSTNGIKILIFSKGY 732  
DB 1081 INNTNNLNLOASKFKLLSINPNKQY 1105

RESULT 14  
C97033  
uncharacterized protein, probably surface-located [imported] - Clostridium acetobutylic  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: C97033  
R:Nolling, J.; Brenner, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C.  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: C97033  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1227 <KUR>  
A:Cross-references: GB:AA01437; PIDN:AAK79054.1; PID:g15023995; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC1080

Query Match 4.7%; Score 179; DB 2; Length 1227;  
Best Local Similarity 18.9%; Pred. No. 0.29;  
Matches 166; Conservative 136; Mismatches 288; Indels 288; Gaps 44;

QY 37 SSTTGDLISFSSSELENTPEYQFQSAIW--SGFIKVKKSDYFTFATSADNHVTMWVDDQF 95  
DB 14 SVTTSVALLSKPAKAYAADNSVKYSISNSNEINVKGE-----VQKEK 58  
QY 96 V--INKASNSKIRLEKRLYQIKIQVORENPTK-----GLDFKLYWTDQNKKEVI 146  
DB 59 VVANKNDSNS--KVSSSENQVS---NKENSNEKVSSEISQINKVNLQVQNNKSVL 114  
QY 147 SSDNLQLPKQKSSNPKK-----RSTSAGTPVDRDNDGIPDSLEVEGYTVDVKNKT 201  
DB 115 AASNVDEVKINSNGVQTSYTAIGETKVKPDTLDIINKAIVDARSAG--TDLSEVEI 172  
QY 202 F-----LSPWISNI-----HEKKG-LTKYK 220  
DB 173 YDIVSQTAETAEAFRNIDGVANVSDYTLGATFVNDANLDSVKNYFYHKRYATVTKFX 232  
QY 221 SSPKWNSTA-----SDPYSDFEKVTGRIDKGVSPARHPLVAA 258  
DB 233 DVATKTSNALKNINNGOGGETDYALEVSGVTGYLDL-----VKNKIYKEQ-----280  
QY 259 YPIVHVDMENIILSKNEDQS--TQNTDSETRISK--NTSTSTRTHTSEVHGNAEVHASFDDI 316  
DB 281 -----NKGRLTITISDSAASTIARINTALDN-----MDA 311  
QY 317 GGSVSAGFSNSNSTVAIDHSL---SLAGERTWAETM---GLNTADTARLANI-----364  
DB 312 GVATLEDYQAI--GANNVPQLHVADYNSLAMPQORGVDSIAIDGINTINTYINNINSVGTE 371  
QY 365 -RYNTGTGAPI-----YVNLPTTSL--VLGKNQTLATIKAKENQLSQILAPNNY---411  
DB 372 DDYINSHAVDSNEGNIDYDILNANIIEKTKAGQDL--TIPEVANVKEVKTLLDFYNHAA 430  
QY 412 -----PSKNLAPIALNAQKFSSTPTIMYNGNFOLEKTKQLRLDQV-----455  
DB 431 AGQTTLDQYKNVDENAOVQDDVAT-----LSMDLXTRDCKTKLALQDKIDSILSNKN 484

```
QY 456 ---YGNATYNFENGRVYRVTGNSWSEVLQIQETIARI-1PNGKOLNVERIAAVNP 510
Db 485 INSGIGNIDDIS-----KJQTEAVDASKLRAVNDIDIKKADKGRDLTIQEIIRDSVKKT 538
QY 511 SDPLETT---KPDMTLKEALKIAF-GFNEPENGLOYQKDIITFDNFDOQTSONIK-- 563
Db 539 IDYINSTGNVSGDGSVDYITIGIDGVTEI--NIEFVNERIKESGITI--TIENIKWV 593
QY 564 ---NOLAEIN--ATNIYTVLD-----KIKLNAXNNILIRKRFHYD 599
Db 594 IEPVOLSEVYRIVTGVTVVYKTLGINNVNDNNIYINAEKKNKDVKIQIDQTRVD 653
QY 600 R--NNI-----AVGADES-----VKEAHEVINSSTEGLLNIDKIRKILSYIVE 645
Db 654 TINNIDVINKIGADVLSDYFNIGITDVQDILDYVADLKIQNYKQVDDIIRVEAK 713
QY 646 IEDTEGLK-----EVINDRYDMLNLSLRQDGKTFIDFKKYNKPLY-----ISNPNYK- 695
Db 714 ISSYBALMRINIGEAVTDDFKALGTDI-----NDGLLYATYDLDQNKYKT 760
QY 696 -----VNVY-AVTKENTIINPSENGDTSTNGI 721
Db 761 ADEVIAVQAQIEIYRALMQIN--LGKATTADYNTLGI 796

RESULT 15
A86827
hypothetical protein yqfG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: A86827
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: A86827
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <STO>
A:Cross-references: GB:AE005176; PID:gi2724625; PIDN:AAK05715.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yqfG

Query Match 4.7%; Score 178.5; DB 2; Length 1072;
Best Local Similarity 19.4%; Pred. No. 0.25;
Matches 160; Conservative 142; Mismatches 342; Indels 179; Gaps 30;

QY 2 VKQENRLNESSSQGLLYFSDLNFOAPMVVTSSTTGDLSIPSELEN----- 52
Db 300 VEQVDVASSESTQDANSASLYPISEASSVTDNLNSISSLDSSISSQTSQSGASSTA 359
QY 53 ---IPSENQYFQSA---IWSGFIKVKSDYEYFATSAHNHVMTWVDQEVINKASNKKI 106
Db 360 EISYDSENGNSJSSNQINSNSKSEKDSQSLGSSMSSEHNSNSNTNETNNSSEI 419
QY 107 R-----LEKRLYQIKIQYQRENPTKEGLDFKL-----YWTDSQKKEVISSDNLQLELK 157
Db 420 TNLPLSPNTEVNSVSDQTSSEASTNSNSISLSPNSISSTSDSESATNSDFSNVAEVA 479
QY 158 QKSSNRKRRKTSAGTVPDRNDGI-----PDSLEVEGYTVYDVKNKRTFLSPWISNIHEK 213
Db 480 NNSLASVNNSSSVLSSTSTADNLGINQSGSDNLTKD--SSEISTSGAFLS---SNQTS 534
QY 214 KGLTKYKS---SPEKWS-----TASDPYDFEKTGTRIDKNVSPRARHPLVAAYPI 261
Db 535 EASTNSNSISLSPNSISSTSVLESTSSNFSNVAEVANNGLASVNNSSSVLSSTSTA 594
QY 262 VHYDM-----ENI-----ILSKNEQSQNTQDSETRTISKNTSTSRHTSEVH 304
Db 595 DNLEINQFGSDNLTKDSSBISTSGAFLSSNQTSSEASSNMSMSSINSPLSLSLTNSESA 654
```

```
QY 305 GNAEVHASFDDIGGSVSAGFSNSNSSTVAIDHSLSL--AGERTWAETMGLTADTARLNA 362
Db 655 TQO-----SNSEATKVDNNST-----HSSNILNSGSDSDSDSDSDSDSSNL-- 699
QY 363 NIRYVNTGAPINVLPTTSLVLGKNQTLATKAKENQLSQ-----ILAPNNTYPSKNLAP 418
Db 700 -----SSPNLETNQITSSKPEVNNISENPVKYSSNSVQENSTDEH 742
QY 419 IALNAOKDFSSPTITWNYNQFLEKTKQLRLDQOVYGNIAATYN---PENGVRVYDTCG 475
Db 743 MSTNPKSSISSPISITSSSQOKESQSN---LLNTTEGINNPITFNNSSSENSAASILT-- 797
QY 476 NMSEVLPOIQETIARIIF-----NGKOL-----NLVERIAAVNPSPDPLETTKPDMT 522
Db 798 SYSNNNSSESSETGCLYISNEAQRDNGSEISHSLPSSNSNNVNSVSIQSOAILLESSKSTN 857
QY 523 LK-EALKIAGFNEP-NGNLOYQGGKDIETFDNFDOQTSONIKNQLAEINATNIYTVLDK 580
Db 858 KRSSLSLIIINSTSHPNEDNQNSSD-----EVKSNNVVESILQUNLSINKTTHMS 909
QY 581 I---KLNAXNNILIRKRFHYDRNNIAGVADSIVKAEHREVINSSTEGLLNIDKIRK 637
Db 910 LTSQKLSVIYTLPSKSKVTNEKNENSTVSEKLIKTPQKN-DESONLQITALDLSFNK 968
QY 638 ILSGYIVEIETDEGLKEVIN-----DRYDMLNLSLRQD-GKTFIDFKKYNKPLY 688
Db 969 -----EVETMEDSKTVPDKVLDNENGRSONKNTSTIAKDNKVKFFKRSEFNSKIL-- 1019
QY 689 ISNPNYKVVYAVTKENTIINPSENGDTSTNGIKKILIFSKKG 731
Db 1020 -----DSDNNILKKTIVLLKKG 1036

Search completed: May 3, 2004, 19:41:13
Job time : 17.0855 secs
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 8.2539 Seconds  
(without alignments)  
4636.784 Million cell updates/sec

Title: US-09-848-909A-8

Perfect score: 3773  
Sequence: 1 EVKQENRLNSESSESSQGLL.....TSTNGIKKILIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3764	99.8	764	1	PAG BACAN
2	235.5	6.2	204	1	YPBI_BACAN
3	185.5	4.9	1630	1	MSPI_PLAFLK
4	185.5	4.9	1639	1	MSPI_PLAFLW
5	175	4.6	1858	1	YM67_YEAST
6	169.5	4.5	1230	1	SMC3_YEAST
7	168	4.5	1957	1	SPOF_SCHPO
8	166.5	4.4	2234	1	WAPA_BACSU
9	164	4.3	1803	1	YJL3_YEAST
10	162.5	4.3	979	1	P115_MYCHR
11	161	4.3	1276	1	BXD_CLOBO
12	159.5	4.2	2869	1	RBP1_PLAFLV
13	159	4.2	1000	1	S155_YEAST
14	159	4.2	1385	1	FAT1_SCHPO
15	158.5	4.2	1167	1	CAGA_HELPJ
16	157	4.2	1790	1	USO1_YEAST
17	156.5	4.1	1208	1	PCP1_SCHPO
18	156	4.1	1928	1	MY81_YEAST
19	156	4.1	2867	1	RBP2_PLAFLV
20	155.5	4.1	2116	1	MIS2_DICDI
21	154.5	4.1	1024	1	RIP3_MOUSE
22	154.5	4.1	1487	1	MDS3_YEAST
23	154	4.1	1037	1	KCC4_YEAST
24	153.5	4.1	678	1	YNC7_YEAST
25	153.5	4.1	1233	1	YFIC_YEAST
26	153	4.1	1288	1	VACB_HELPJ
27	152.5	4.0	1116	1	YK54_AQUAE
28	152	4.0	1029	1	RIP3_RAT
29	152	4.0	1420	1	SRB9_YEAST
30	151	4.0	1882	1	Y468_MYCPN
31	149	3.9	918	1	YHJB_CABEL
32	149	3.9	1577	1	HLVA_PROMI
33	148.5	3.9	1271	1	Y338_MYCGE

34 148 3.9 1250 1 BXE\_CLOBO  
35 147.5 3.9 1225 1 Y305\_MYCGE  
36 147.5 3.9 1450 1 N159\_YEAST  
37 147.5 3.9 1726 1 MSPI\_PLAFLP  
38 147.5 3.9 2376 1 TAO3\_YEAST  
39 146.5 3.9 730 1 GLN3\_YEAST  
40 146.5 3.9 1091 1 CIC2\_RAT  
41 146.5 3.9 1162 1 BXEN\_CLOBU  
42 146.5 3.9 1388 1 SIR4\_YEAST  
43 146.5 3.9 1397 1 CID\_DROME  
44 146.5 3.9 2339 1 RPI1\_PLAFA  
45 146 3.9 719 1 YMA1\_YEAST

Q00496 clostridium  
P47551 mycoplasma  
P40477 saccharomyc  
P50495 plasmodium  
P40468 saccharomyc  
P18494 saccharomyc  
P54290 rattus norv  
Q06366 clostridium  
P11978 saccharomyc  
P27625 plasmodium  
Q03213 saccharomyc

#### ALIGNMENTS

RESULT 1  
PAG\_BACAN  
ID PAG\_BACAN STANDARD; PRT; 764 AA.  
AC P13423; Q9FSR7; Q9KH69; Q9RQU2;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins  
DE translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].  
GN PAGA OR PAG OR PXO1-110.  
OS Bacillus anthracis.  
OG Plasmid pXO1.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89172073; PubMed=3148491;  
RA Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppia S.H.,  
RA Schmidt J.J.;  
RT "Sequence and analysis of the DNA encoding protective antigen of  
RT Bacillus anthracis.";  
RL Gene 69:287-300(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=28, 33, BA1024, and BA1035;  
RX MEDLINE=99214082; PubMed=10197996;  
RA Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;  
RT "Genetic diversity in the protective antigen gene of Bacillus  
RT anthracis.";  
RL J. Bacteriol. 181:2358-2362(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V770-NP1-R / ATCC 14185;  
RX MEDLINE=20359347; PubMed=1089854;  
RA Cohen S., Mendelson I., Altbaum Z., Kobiler D., Elhanany E., Bino T.,  
RA Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,  
RA Kronman C., Velan B., Shaffer A.;  
RT "Attenuated nontoxicogenic and nonencapsulated recombinant Bacillus  
RT anthracis spore vaccines protect against anthrax.";  
RL Infect. Immun. 68:4549-4558(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Stearne;  
RX MEDLINE=99445483; PubMed=10515943;  
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,  
RA Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,  
RA Martinez Y., Rieke D., Svensson R., Jackson P.J.;  
RT "Sequence and organization of pXO1, the large Bacillus anthracis  
RT plasmid harboring the Anthrax toxin genes.";  
RL J. Bacteriol. 181:6509-6515(1999).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=99445483; PubMed=10515943;  
RX MEDLINE=91332080; PubMed=1651334;  
RA Singh Y., Kilmpel K.R., Quinn C.P., Chaudhary V.K., Leppia S.H.;  
RT "The carboxyl-terminal end of protective antigen is required for  
RT receptor binding and anthrax toxin activity.";

J. Biol. Chem. 266:15493-15497(1991).  
[6]  
CHARACTERIZATION.  
RC STRAIN=Stearne;  
RX MEDLINE=94327640; PubMed=8051159;  
RA Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;  
RT "Anthrax protective antigen forms oligomers during intoxication of  
mammalian cells."; J. Biol. Chem. 269:20607-20612(1994).  
[7]  
CHARACTERIZATION.  
RN STRAIN=Stearne;  
RX MEDLINE=21129592; PubMed=11207581;  
RA Beaugregard K.E., Collier R.J., Swanson J.A.;  
RT "Proteolytic activation of receptor-bound anthrax protective antigen  
on macrophages promotes its internalization."; Cell. Microbiol. 2:251-258(2000).  
[8]  
TOXIN REGULATION.  
RN STRAIN=Weybridge;  
RX MEDLINE=94131936; PubMed=8300513;  
RA Koehler T.M., Dai Z., Kaufman-Yarbray M.;  
RT "Regulation of the Bacillus anthracis protective antigen gene: CO2 and  
a trans-acting element activate transcription from one of two  
promoters."; J. Bacteriol. 176:586-595(1994).  
[9]  
MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.  
RN STRAIN=Stearne;  
RC STRAIN=Stearne;  
RX MEDLINE=95050722; PubMed=7961869;  
RA Singh Y., Klimpel K.R., Arora N., Sharma M., Leppla S.H.;  
RT "The chymotrypsin-sensitive site, PFD315, in anthrax toxin protective  
antigen is required for translocation of lethal factor."; J. Biol. Chem. 269:29039-29046(1994).  
[10]  
MUTAGENESIS OF DOMAIN 4 LOOPS.  
RN STRAIN=Stearne;  
RC STRAIN=Stearne;  
RX MEDLINE=99185012; PubMed=10085028;  
RA Varughese M., Teixeira A.V., Liu S., Leppla S.H.;  
RT "Identification of a receptor-binding region within domain 4 of the  
protective antigen component of anthrax toxin."; Infect. Immun. 67:1860-1865(1999).  
[11]  
MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.  
RN STRAIN=Stearne;  
RC STRAIN=Stearne;  
RX MEDLINE=21092804; PubMed=11178978;  
RA Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;  
RT "Trp 346 and Leu 352 residues in protective antigen are required for  
the expression of anthrax lethal toxin activity."; Biochem. Biophys. Res. Commun. 281:186-192(2001).  
[12]  
MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.  
RN STRAIN=Stearne;  
RC STRAIN=Stearne;  
RX MEDLINE=21438996; PubMed=11554763;  
RA Ahuja N., Kumar P., Bhatnagar R.;  
RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are  
required for oligomerization of anthrax protective antigen."; Biochem. Biophys. Res. Commun. 287:542-549(2001).  
[13]  
MUTAGENESIS OF PRO-289.  
RN STRAIN=Stearne;  
RC STRAIN=Stearne;  
RX MEDLINE=21255689; PubMed=11356563;  
RA Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;  
RT "Role of residues constituting the 2beta1 strand of domain II in the  
biological activity of anthrax protective antigen."; FEBS Microbiol. Lett. 199:27-31(2001).  
[14]  
MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.  
RX MEDLINE=21125576; PubMed=11222612;  
RA Mogridge J., Mourez M., Collier R.J.;  
RT "Involvement of domain 3 in oligomerization by the protective antigen  
moiety of anthrax toxin."; J. Bacteriol. 183:2111-2116(2001).  
[15]

MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.  
RX MEDLINE=21269403; PubMed=11113126;  
RA Sellman B.R., Nassi S., Collier R.J.;  
RT "Point mutations in anthrax protective antigen that block  
translocation."; J. Biol. Chem. 276:8371-8376(2001).  
[16]  
MUTAGENESIS OF PRO-213; LEU-216; PHE-231; LEU-232; PRO-234; ILE-236;  
RP ILE-239; TRP-255 AND PHE-265.  
RC STRAIN=Stearne;  
RX MEDLINE=22112896; PubMed=12117959;  
RA Chauhan V., Bhatnagar R.;  
RT "Identification of amino acid residues of anthrax protective antigen  
involved in binding with lethal factor."; Infect. Immun. 70:4477-4484(2002).  
[17]  
X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RX MEDLINE=97192099; PubMed=9039918;  
RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;  
RT "Crystal structure of the anthrax toxin protective antigen."; Nature 385:833-838(1997).  
[18]  
REVIEW.  
RX MEDLINE=21428689; PubMed=11544370;  
RA Mock M., Fouet A.;  
RT "Anthrax"; Annu. Rev. Microbiol. 55:647-671(2001).  
CC -!- FUNCTION: One of the three proteins composing the anthrax toxin,  
the agent which infects many mammalian species and that may cause  
death. PA binds to a receptor (ATR) in sensitive eukaryotic  
cells, thereby facilitating the translocation of the enzymatic  
toxin components, edema factor and lethal factor, across the  
target cell membrane. PA associated with LF causes death when  
injected, PA associated with EF produces edema. PA induces  
immunity to infection with anthrax.  
CC -!- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a  
protective antigen (PA), a lethal factor (LF) and an edema factor  
(EF). None of these is toxic by itself. PA+LF forms the lethal  
toxin (Leth); PA+EF forms the edema toxin (EdTX). PA-63 forms  
heptamers and this oligomerization is required for LF or EF  
binding. Once activated, at low pH, the heptamer undergoes  
conformational changes and converts from prepore to pore inserted  
in the membrane, forming cation-selective channels.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DOMAIN: The molecule is folded into four functional domains. Each  
domain is required for a particular step in the toxicity process.  
Domain 1 contains two calcium ions and the proteolytic activation  
site. Cleavage of the PA monomer releases the subdomain 1a, which  
is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is  
part of the remaining 63-kDa fragment (PA63) and contains the  
binding sites for LF and EF. Domain 2 is a beta-barrel core  
containing a large flexible loop that has been implicated in  
membrane insertion and pore formation. There is a chymotrypsin  
cleavage site in this loop that is required for toxicity. Domain 3  
has a hydrophobic patch thought to be involved in protein-protein  
interactions. Domain 4 appears to be a separate domain and shows  
limited contact with the other three domains: it would swing out  
of the way during membrane insertion. It is required for binding  
to the receptor; the small loop is involved in receptor  
recognition.  
CC -!- PTM: Proteolytic activation by furin or a furin-like protease  
cleaves the protein in two parts, PA-20 and PA-63; the latter is  
the mature protein. The cleavage occurs at the cell surface and  
probably in the serum of infected animals as well; both native and  
cleaved PA are able to bind to the cell receptor. The release of  
PA20 from the remaining receptor-bound PA63 exposes the binding  
site for EF and LF, and promotes oligomerization and  
internalization of the protein.  
CC -!- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were  
performed that showed that the residues present in the small loop  
of domain 4, and not the ones in the large loop, are involved in  
receptor recognition.  
CC -!- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

MEDLINE=89172073; PubMed=3148491;  
RA Welkos S.L., Lowe J.R., Eden-McCuchan F., Vodkin M., Leppla S.H.,  
RA Schmidt J.J.;  
RA "Sequence and analysis of the DNA encoding protective antigen of  
RT Bacillus anthracis";  
RL Gene 69:287-300(1988).  
RW [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Stemte;  
RC MEDLINE=99445483; PubMed=10515943;  
RX Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hall K.K.,  
RA Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,  
RA Martinez Y., Rieke D., Svensson R., Jackson P.J.;  
RA "Sequence and organization of pXOI, the large Bacillus anthracis  
RT plasmid harboring the Anthrax toxin genes.";  
RL J. bacteriol. 181:6509-6515(1999).  
RC -!- SUBCELLULAR LOCATION: Membrane-bound (Potential).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M22589; AAA22636.1; -;  
DR EMBL; AF065404; AAD32415.1; -;  
DR PIR; G59104; G59104.  
DR PIR; I39933; I39933.  
DR HSP; P13423; IACC.  
DR Hypothetical protein; Plasmid; Transmembrane.  
FT TRANSMEM 162 182  
FT CONFLICT 93 93  
FT CONFLICT 184 204  
FT  
FT  
SQ SEQUENCE 204 AA; 23029 MW; E1657B23AE4273FD CRC64;  
  
Query Match 6.2%; Score 235.5; DB 1; Length 204;  
Best Local Similarity 34.6%; Pred. No. 2.5e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;  
  
QY 587 MNTLLDKRPHYDNNIATGADSVKAEHREVINSSTEGLLNIDKDIRKILSGYVEI 646  
Db 1 MNTLVDRP-VHYDNNGNIVGDSYLKNAKYQLINWSGVSGLNLDVEDVQALSGYMLQI 59  
  
QY 647 EDTE-----GLKEVINDRYDMLNISLRODGKTFDFKYNKDKLPYISNP 693  
Db 60 KPSPNHLTNSPVTITLAKDQSGVGLYRVL-----DGTGFLDNKFDENWRLV-DPG 112  
  
QY 694 YKNVYAVTKEN-TIINPSENGDTSTNGIKILIFSXKGYEI 734  
Db 113 DDVYVYAVTKEDFNVAVTRDENGNTA-NKIKNTLVLSKIKVEI 153  
  
RESULT 3  
MSPI\_PLAFK  
ID MSPI\_PLAFK STANDARD; PRT; 1630 AA.  
AC P04932;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)  
DE (PMMSA) (P190).  
GN MSP-1.  
OS Plasmodium falciparum (isolate K1 / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86136024; PubMed=3004972;  
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,  
RA Stunnenberg H., Bujard H.;

RT Polymorphism of the precursor for the major surface antigens of  
RT Plasmodium falciparum merozoites: studies at the genetic level.;  
RL EMBO J. 4:3823-3829 (1985).  
RN [2]  
RP REVISIONS, SEQUENCE FROM N.A.  
RA Pan W., Tolle R., Bujard H.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
CC (potential).  
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42  
CC kDa and 19 kDa antigens which are the major surface antigens of  
CC merozoites. The maturation take place during schizont.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X03371; CA227070.1; -  
DR InterPro; IPR006209; EGF-like.  
DR Pfam; PF00008; EGF; 1.  
DR Malaria; Merozoite; Polyprotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor.  
FT SIGNAL 1 19  
FT CHAIN 20 1630  
FT DOMAIN 67 84  
FT TRANSMEM 1614 1630  
FT CARBOHYD 97 97  
FT CARBOHYD 259 259  
FT CARBOHYD 755 755  
FT CARBOHYD 759 759  
FT CARBOHYD 774 774  
FT CARBOHYD 835 835  
FT CARBOHYD 911 911  
FT CARBOHYD 955 955  
FT CARBOHYD 1049 1049  
FT CARBOHYD 1156 1156  
FT CARBOHYD 1165 1165  
FT CARBOHYD 1436 1436  
FT CARBOHYD 1517 1517  
FT SEQUENCE 1630 AA; 187289 MW; ADBD8C3CE0A46322 CRC64;  
Query Match 4.9%; Score 185.5; DB 1; Length 1630;  
Best Local Similarity 20.5%; Pred. No. 0.12;  
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;  
QY 5 ENEL-LNESESSQGLGYFDFNFQAPMVVTSSTGDLSPSSSELENIPSENYQFOSA 63  
DB 918 ENILSGKNKVIYQELIGQKSB-NF-----YEKILKDSITFYNE 956  
QY 64 IWSGFIKVKSDBYTATFATSADNHVTWVDQEVINKASNS-----NKIBLKGRLY- 114  
DB 957 SFTNFVKSADD-----INSLNDESKRKKLEEDINKLKTQLSFDLYNKYKLERLPD 1011  
QY 115 -----QIKIYOQRENPTKGLDFKLYWTDSONKKEVISDNLQLPKOKSNSRKKHS 168  
DB 1012 KKTIVGKYNQIKKLTLLKEQLSKU--NSLNPKHVL--QNPSPFNKKKEAEATEN 1067  
QY 169 TSAGPTVPDRDNGIDPSLEVEGYTDVKNKRTFLSPWISNTHKKGLTKY---KSSP-- 223  
DB 1068 T-----LENTKILLKHY-----KGLVKYNGESSPLK 1094  
QY 224 ---EKWSTASDPYSDPF-----KVTRIDKNVSPAR-----HPLVAAPYIVHVD 265  
DB 1095 TLSEESIQTEDNYASLENFKVLKLEGLKNDLNLEKKKLSYLSGLHLLIA-----E 1147  
QY 266 MENILSKNEDQSTQNTDSETRITSKNTSRTSTSEVHGNAVHASFDFDGSVSAGFS 325  
DB 1148 LKEVIKKNQ---YTGNSPSENNT-----DVNNALESYKFLPEGTDVATVVS 1191

QY 326 NSNSTVAIDHSLAGERTWAETMGLNTADTARLNANIRYVNTGTAPINVLPTT--- 381  
DB 1192 EGSQDTLEOSQPKPASTHVGAE---NTIITSQ-NVDEVDVVIIVIFGSEEDYDDL 1247  
QY 382 -SLVLGKNOTLATIKAKENQLSQAIPANNYPSKNVLAFA---LNAQKDFSTPTTMVYN 437  
DB 1248 GQVVTGEAVTPSVI---DNILSKI---ENEVEVLYLKLPLAGVYRSLEKQLENVNTFVN 1301  
QY 438 -----QFLELEKTKQLRLDQV-YGNIAI-----YNFENGRVVRDVTGSNWSEVL 481  
DB 1302 VVDILNSRNKRNKFNQV-LESDLIPYKDLTSSVYVVDKPYKFLNKKRDKFLSYNYIK 1360  
QY 482 POIQETTARIIFNGKDLNLVERRIAAVNPSDPL-----ETTKPDM-TLKALKIAPGF 533  
DB 1361 DSDI-----TDINEFA-----NDVLGYVKILSEKYSKSDLSIKKIYINDKQGE 1401  
QY 534 NEPN-----GNLQVQGGKITE-----PDPNFDQOTSQNIKNQALNAATNIYT 576  
DB 1402 NEKYLPLNNIETLYKTVNDKIDLFVHLEAKVLYNYTEK---SNVEVKIKELN--YLKT 1456  
QY 577 VLDKIKLNARMN--ILIRDKRFHYDRNN-----IAGV-ADESVVKEAHREVINSTEGLL 628  
DB 1457 IQKLADFKKNNFVGIADLSTDYNNHLLTKFLSTGMVFENLAKTLVLSNLDGNLQGM 1515  
QY 629 LNTDKD--TRKILSGYVIEIETEGLEKVINDRYDMLNLSLRQDGKTFIDFKYNDKLP 686  
DB 1516 LNTSQHCQVKK-----QCPQNSGCFRHLDE-----RECKCLLNTYKQEGDKC- 1557  
QY 687 LYISNPNYKVNVAVTAKNTIINPSNG-----DTSTNGIKKI 724  
DB 1558 --VENFNPTCN-----EN-----NGGCDADAKTEEDSGSNG-KKI 1590  
RESULT 4  
MSPI PLAPW STANDARD; PRT; 1639 AA.  
AC P04933;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)  
DE (PMMSA) (P195).  
GN MSP-1.  
OS Plasmodium falciparum (isolate Wellcome).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5948;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86014355; PubMed=2995820;  
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,  
RA Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,  
RA Freeman R.R.;  
RT "Primary structure of the precursor to the three major surface  
RT antigens of Plasmodium falciparum merozoites.";  
RL Nature 317:270-273 (1985).  
RL [2]  
RP REVISIONS.  
RA Holder A.A.;  
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
CC (potential).  
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42  
CC kDa and 19 kDa antigens which are the major surface antigens of  
CC merozoites. The maturation take place during schizont.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC



EMBL; X02919; CAA26676.1; --  
DR PIR; A24594; A24594.  
DR PIR; S05603; S05603.  
DR PDB; 1CEJ; 28-MAY-99.  
DR InterPro; IPR006209; EGF\_like.  
DR Pfam; PF0008; EGF; 1.  
KW Malaria; Mercoite; Polyprotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor; 3D-structure.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1639 MERZOITE SURFACE PROTEIN 1.  
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 768 768 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 783 783 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 844 844 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 964 964 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1165 1165 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1526 1526 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1639 AA; 187618 MW; 2C255B6616C87F6E CRC64;  
Query Match 4.9%; Score 185.5; DB 1; Length 1639;  
Best Local Similarity 20.5%; Pred. No. 0.12;  
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;  
QY 5 ENRL-LNESESSQGLLYYFSDLPQAPMVVTSSTTGLSPSELENISENQYFOSA 63  
DB 927 ENILSGKKNYQELIGQSSSE-NF-----YKILKSDSTFVNE 965  
QY 64 IWSGFIKVKSDYEYFATSADNHVTWVDQVINKASNS-----NKIRLEKGRLY- 114  
DB 966 SFTNFVKSADP-----INSLNDESKRKLLEIDINKLTKLQSLDLYNKVLELRF 1020  
QY 115 -----QIKIOYQRENPTKGLDFKLYMTDSQNKKEVSSDNLQPELKXSSNRKRS 168  
DB 1021 KKTGVCKYKQKJLLKLEQESKL--NSLNPHVLI--QNFVFFKKEAEIAETEN 1076  
QY 169 TSAGPTVPDRNDGIPDSLEVEGYTVVKNKFTFLSPMISNHEKKGATKY--KSSP-- 223  
DB 1077 T-----LENTKILLKHY-----KGLVKNYNGESSPLK 1103  
QY 224 ---EKWSTASDPYSDPE-----KYTGRIKDNVSPPEAR-----HPLVAAPYIVHD 265  
DB 1104 TLESISQIEDNYASLENFKVLSLEGKLNLEKKKLSYLSGLHLLIA-----E 1156  
QY 266 MENILSKNEDQSTQNTDSETRTSKNTSTSTRTSEVHGNAEVSASFIDIGSVSAGFS 325  
DB 1157 LKEVIKKNK---VTGNSPSENNT-----DVNNALSYKKFLPBGTDVATVS 1200  
QY 326 NNSSTVAIDHLSLAGERTWAGTGLNLTADTARLNANIRVYNTGTATYINVLPTT---- 381  
DB 1201 EGSQDTLEQSPKPKASTHVGAES---NTITTSQ-NVDDEVDVIVIFIGSEEDYDDL 1256  
QY 382 -SLVLGKNOTLATIRAKENQLSQILAPNNYPSKNLAPIA---LNAQKDFGSPITMNYN 437  
DB 1257 GQVVTGEAVTPSVI---DNILSKI---ENEYEVLYKPLAGVYRSLKQLENNVMTFN 1310  
QY 438 -----QFLELEKTKQLRDTDOV-YGNIAF-----YFNGRVRVDTGNSWEVL 481  
DB 1311 VKDILNSRENKFNKVN-LESGLPYKDLTSSNVVVKDPYKFLNKEKRDKFLSSYNIK 1369  
QY 482 POIQETARIIIPNGKDLNVERRIAANVPSDPL-----ETTKPDM-TLKEALKIAPGF 533  
DB 1370 DSID-----TDINFA-----NDVLGYKKILSEKYSKLDLSIKYINDKQGE 1410  
QY 534 NBNP-----ONLYQVQKDIPE-----FDFNFQDQTSQNKQLAELNATNIYT 576  
DB 1411 NEKLYPFLNIIETLYKTVNDKIDLFVHLEAKVNLNTEYK---SNVEVKIKELN--YLKT 1465

```
Db 460 EQQENDEPEKDDIIRSSIDKFNHGNNKSEVSENVLENEDTPAIVERENQ-IND---VE 515
Qy 191 GTVDVVKNRKPTL-----SWISNIHR--KKGLTKYKSS-----PEKNSTASD----- 231
Db 516 GYDVTGKSVESDLHESP--DNLYDLAARAMLQFOQSRNSNCPQKEQVSSYLGHSNGS 573
Qy 232 -----PYSDFEKTGRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQ 277
Db 574 NLSGRSLDESBEQILKDF--TGENNNKLT-DGDLSSS---VEIEVEKV---SEKK 622
Qy 278 STQNTDSEPTISKNSTSRTHTSVHGNAEVHAFDIDGGSVAGFSNNSSTVAIDHS 337
Db 623 LDGSTEKELVPLSTDTTIN-----NSSLGNEDSIYSLDDA 658
Qy 338 LSLAGERTWAETMGLNTA-----DTARLNANIRVYVNTGTAPIYVNL 378
Db 659 DAISENLTVPLMEIKTKYEVWISVSYSSTVEDNTVAMPPOVEY-----TSFPMND- 713
Qy 379 PTTSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIALNAQKDFSSPTITM----- 434
Db 714 PFNSL-----ND---DYEKKHLLKSTLAA-----LAPFTKDAEFVAGVTKSLT 758
Qy 435 ----NYNQFLELEKTKOLALDQVYGNIAFYNFENGRVVRTGSGNWSEVLPOIQETAR 490
Db 759 STSGHTNIHTSKETKQVS--DUDESTENVTFENVTG-----DENKNQKNPFGVANSTDK 813
Qy 491 IIFNGKDLNLVERRIAANPSPDLETTKPDMLTKEALKIAFGNFENGLQVQGDITFEF 550
Db 814 STEDNTD---EKYFSAINTVN---VTGDSCEDIETASVVE---NLRYCEKDMNEA 862
Qy 551 DFNF-DQOTSON---IKNQLA-----ELNATNIYTVLDKTKLNAKNLILRDKRFHYD 599
Db 863 EMSGGDECVQNDGSKTQISFTDPSDFPNQESNDNTEFSSTK-----YK 907
Qy 600 RNNIAGADESVVKEAHR-EVIN-----SSTEGLLNIDKDIRKI 638
Db 908 VRNSDLEDDSLAKELTKAEVVKLDEESSEDSVEQDYADPEPQNDGSENIYKGTK- 966
Qy 639 LSGYVIEIETEGLEKVINDRYDMLNLSLRQDKTFLDFKYNKDLKPLVISPKNVKNV 698
Db 967 -----DTLGIVPEPEKVN-----KVHEETLFEANVSSSVNV 999
Qy 699 YAVTKENTIIN 709
Db 1000 QNKDMHTDVIN 1010

RESULT 6
SMC3_YEAST STANDARD; PRT; 1230 AA.
AC P47037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Structural maintenance of chromosome 3 (DA-box protein SMC3).
GN SMC3 OR YJL074C OR J1049.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RX MEDLINE=97474309; PubMed=9335333;
RA Michaelis C., Ciosk R., Nasmyth K.;
RT "Cohesins: chromosomal proteins that prevent premature separation of
sister chromatids";
RL Cell 91:35-45 (1997).
[2]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=96208490; PubMed=8641269;
RA Galibert F., Alexandraki D., Baur A., Boles E., Chalwatzis N.,
```

```
Chuat J.-C., Coster F., Cziepluch C., De Haan M., Domdey H.,
Durand P., Enlian K.-D., Gatus M., Goffeau A., Grivell L.A.,
Huangmann A., Herbert C.J., Heumann K., Hilger F., Hollenberg C.P.,
Huang M.-E., Jacq C., Jauniaux J.-C., Katsoulou C., Kirchthar L.,
Kleine K., Kordes E., Koetter P., Liehl S., Louis E.J., Manus V.,
Mewes H.-W., Miosga T., Obermaier B., Perea J., Pohl T.M.,
Portetelle D., Pujoil A., Fumelle B., Ramezani Rad M., Rasmussen S.W.,
Rose M., Rossau R., Schaaff-Gerstenschlaeger I., Smits P.H.M.,
Scarcez T., Soriano N., Tovan D., Tzermia M., Van Broekhoven A.,
Vanderbol M., Wedler D.H., Von Wettstein D., Wambutt R., Zagulski M.,
Zollner A., Karpfinger-Hartl L.;
"Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
X.";
EMBO J. 15:2031-2049 (1996).
[3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1678;
RX Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[4]
RN IDENTIFICATION IN A COHESIN COMPLEX WITH SMC1, IRR1 AND MCD1, AND
INTERACTION OF THE COHESIN COMPLEX WITH SCC2.
MEDLINE=99145468; PubMed=990856;
RX Toth A., Ciosk R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;
"Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to
establish cohesion between sister chromatids during DNA replication.";
Genes Dev. 13:320-333 (1999).
[5]
RN IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3, MCD1 AND IRR1, AND
STRUCTURE.
MEDLINE=21980169; PubMed=11983169;
RX Haering C.H., Loewe J., Hochwagen A., Nasmyth K.;
"Molecular architecture of SMC proteins and the yeast cohesin
complex.";
Mol. Cell 9:773-788 (2002).
CC -!- FUNCTION: Involved in chromosome cohesion during cell cycle and in
DNA repair. Central component of cohesin complex. The cohesin
complex is required for the cohesion of sister chromatids after
DNA replication. The cohesin complex apparently forms a large
proteinaceous ring within which sister chromatids can be trapped.
At anaphase, the complex is cleaved and dissociates from
chromatin, allowing sister chromatids to segregate.
CC -!- SUBUNIT: Cohesin complexes are composed of the SMC1 and SMC3
heterodimer attached via their hinge domain, MCD1/SCC1 which link
them, and IRR1/SCC3, which interacts with MCD1. The cohesin
complex also interacts with SCC2, which is required for its
association with chromosomes.
CC -!- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin.
Before prophase it is scattered along chromosome arms. At
anaphase, the MCD1 subunit of the cohesin complex is cleaved,
leading to the dissociation of the complex from chromosomes,
allowing chromosome separation.
CC -!- DOMAIN: The flexible hinge domain, which separates the large
intramolecular coiled coil regions, allows the heterotypic
interaction with the corresponding domain of SMC1, forming a V-
shaped heterodimer. The two heads of the heterodimer are then
connected by different ends of the cleavable MCD1 protein, forming
a ring structure (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. SMC3 subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; Y14278; CAA74655.1; -.
CC EMBL; Z49349; CAA89366.1; -.
CC EMBL; X88851; CAA61313.1; -.
CC F01; S56850; S56850.
CC GenOnline; 141688; -.
DR DR
```

DR SGD; S0003610; SMC3.  
DR GO; GO:0007151; P:sporulation (sensu Saccharomycetes); IMP.  
DR GO; GO:0007130; P:synaptonemal complex formation; IMP.  
DR InterPro; IPR003439; ABC transporter.  
DR InterPro; IPR003405; SMC\_C.  
DR InterPro; IPR003395; SMC\_N.  
DR Pfam; PF02463; SMC\_C; 1.  
DR Pfam; PF02463; SMC\_N; 1.  
KW Mitosis; Cell cycle; Chromosome partition; ATP-binding; Coiled coil;  
KW Nuclear protein.  
FT NP\_BIND 32 39 ATP (POTENTIAL).  
FT DOMAIN 172 482 COILED COIL (POTENTIAL).  
FT DOMAIN 483 684 FLEXIBLE HINGE.  
FT DOMAIN 685 1041 COILED COIL (POTENTIAL).  
FT DOMAIN 1126 1161 ALA/ASP-RICH (DA-BOX).  
SQ SEQUENCE 1230 AA; 141336 MW; B152D88F7780341F CRC64;

Query Match 4.5%; Score 169.5; DB 1; Length 1230;  
Best Local Similarity 19.8%; Pred. No. 0.5; Mismatches 355; Indels 247; Gaps 38;  
Matches 182; Conservative 136;

QY 1 EYKQENRLNSESSESGLLGYFSDLNFPQAPMVVT--SSTGTGDSLSPSSSELENIPSENQ 58  
Db EMEQERKELEKYNELERNKIYQFTLYDRELNEVINQMERLGDYN-----NTVYSSEQ 262  
QY 59 YFQ-----SAWSGFIKVKSDYETFATSADNHVTMWVDDOEVINK--- 99  
Db YIQELDKREDMDIQVSKLSSI-EASLKIKATDLOQAKLRESEISQKLTNNVVKIDVQ 321  
QY 100 ---ASNSNIRLEKGLYQIK-IQYQRENPTKGLDFKLYWDSONKKEVISSDNLQPE 155  
Db 322 QQIESNEEQRLNDSATLKEIKSIIEQKQKLSKILPRYQELT---KEEAMY--KLQLAS 375  
QY 156 LKQSNRSKRRTSAGTPVDRDNDGIPDSLEVEGYVDVKNKRTFLSPWISNTHK-KK 214  
Db 376 LQKQKRDILKGEYARFKSKQERDTHISEIE-----ELKSS-----IQNLNELES 422  
QY 215 GLTKVSSPKSTWSDPSYDPKVTGRIDKNVSPPEARPLVAAPVHVDMENIILSKN 274  
Db 423 QLQMDRTSLRKQYSAID-----EEIELDLSINGPTKQL-----EFDSELHLKQK 471  
QY 275 EDGSTQNTDSETHIISKNTSTSTHSTSEVHGNAEVAHFFDGGVSGAGFSNSSTVAI 334  
Db 472 LSESLDTRKELWEKQKLTVEITLLSDVNQNR-----NVNETMSRSLANGIINVKRI 525  
QY 335 DHSISLAGE-----RTWAEITMG-----LNTADTARLNANIRVYNTG 370  
Db 526 TEKLKSPESVFGTGLIKNDKYKTCAEVIGNSLFIHVDTETATLNNELYRMKG 585  
QY 371 TAPIYVLPITSLVLGKNQTLATIKAKENQLSQILAPNNYPSKNLAPIALNAQKDFST 430  
Db 586 GRVTF--IPLNRLSLDSVDFPSNTTQIQFTPLIKKIYEPFEKA-----VGVFVGKT 638  
QY 431 PITMAYNQFLEKTKQLR---LDTQV-YGNLAT--YNFENGRVVDVTSKNWSEVLPOI 484  
Db 639 IVVKDLQGGKLAKKHKLNAITLDGDRADKRGVLTGGYLDQHKRTLESILKNLNSRSOH 598  
QY 485 QETTARIIFNGKOLNVERIRAAVNPSS-----DPLETTKPD-MTLK 524  
Db 699 KKILEELDFVRELNDIDTKIDQVGNIRKVSNDRESVLNTIEVYRTSLNTKKNKXILE 758  
QY 525 EALK-IAFGNPNNGNIQYQKQITDFN-----FDQOTSONKNQIA-----E 568  
Db 759 ESLNAITLKLEKLTNRTFAQEKLTNFENDLQLEFDSLSKEERLESITKEISAHNK 818  
QY 569 LNATN-----IYTVLDIKLNKKNILIRDRKFRHYDRNNIYAVG-----A 607  
Db 819 LNITSDALEGTTTID--SLNAALESKLIQENDLESKMGSEVGDATIFGQDELKELOLE 876  
QY 608 DESVVKRAHREVINSST-----EGLLI-----NIDKDKIRKI 638  
Db 877 KESVEKQEHENAVELGTQVREIBESLIAETNNKKLEKANNQOQLLKLKLDNFQKSEVT 936

QY 639 LSGYVIVEIEDTEGLKE-----VINDRYDM-----LN-----TSSLRODGK 673  
Db 937 MIKKTTLVTRBELQORIREIGLLPEDALPEDAITSQDQLRLNDMTMTSEGLKNVNR 996  
QY 674 -TFIDFKYNDK-----LPLYSNPYKNVYAVTAKENTIIINPSENG 714  
Db 997 RAFENFKFNERKDLAERASELDESKDSIQDLVLKQKQNAV-----DSIFQKYSNF 1052  
QY 715 DT-----STNGIKKILIFSK 729  
Db 1053 EAVFERLVPRGTAKLIHRK 1072

RESULT 7  
SPOF SCHPO  
ID SPOF SCHPO STANDARD; PRT; 1957 AA.  
AC Q10411; Q9USE9;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Sporulation-specific protein 15.  
GN SPO15 OR SPAC1F3.06C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
ON NCBI\_TaxID=4896;  
RX [1]  
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.  
RX MEDLINE=20107136; PubMed=10639340; Shimoda C.;  
RA Ikemoto S., Nakamura T., Kubo M., Shimoda C.;  
RT "S. pombe sporulation-specific coiled-coil protein Spo15p is localized  
to the spindle pole body and essential for its modification";  
RL J. Cell Sci. 113:545-554 (2000).  
RX [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouras J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares K., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Voicakart G., Aert R., Robben J., Grymonprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wadler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe";  
RL Nature 415:871-880 (2002).  
RX [3]  
RP SEQUENCE OF 705-871 FROM N.A.  
RC STRAIN=968 h90;  
RX MEDLINE=20223868; PubMed=10759889;  
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,  
RA Hiraoka Y.;  
RT "Large-scale screening of intracellular protein localization in living  
fission yeast cells by the use of a GFP-fusion genomic DNA library";  
RL Genes Cells 5:169-190 (2000).

CC -!- FUNCTION: Has a role in the initiation of spore membrane  
CC formation.  
CC -!- SUBUNIT: Monomer.  
CC -!- SUBCELLULAR LOCATION: Spindle pole body.  
CC -!- SIMILARITY: Belongs to the Mpc70 family.  
CC  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC  
CC EMBL: Z70690; CAA94624.1; -  
CC EMBL: AB027811; BAA87115.1; -  
CC PIR: T38077; T38077.  
CC GeneDB: Spombe; SPAC1F3.06c; -  
CC Sporulation; Coiled coil.  
CC KW DOMAIN 199 785 COILED COIL (POTENTIAL).  
CC FT DOMAIN 804 1235 COILED COIL (POTENTIAL).  
CC FT DOMAIN 1320 1471 COILED COIL (POTENTIAL).  
CC FT DOMAIN 1481 1723 COILED COIL (POTENTIAL).  
CC SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;  
Query Match 4.5%; Score 168; DB 1; Length 1957;  
Best Local Similarity 20.5%; Pred. NO. 1.1;  
Matches 178; Conservative 132; Mismatches 327; Indels 232; Gaps 40;  
QY 2 VKQENRLLESE-----SSQGLIGYYSDFNFO---APMVVTSSTTGDLSIPSE 49  
DB 129 VTQKLNLELKVQRKLALEHENGILSLQSSNKKDKNTSSVTTLTSEEDSVFQKK 188  
QY 50 LENIPSENYQFSAIMS---GFIVK-----KKSDEYTFATSDADNHVTVWVDQEVINKAS 101  
DB 189 LTNMESNFSKQSEAYDLSQLLTVTETKDKKDYKEDVSSIKASLAEEQASNKL 248  
QY 102 NSKIRLEK---GRLVQIKIYOENPTE---KGLDFKLYWTDSONKKEVISEDNLQPE 155  
DB 249 RGEQERLEKLLVSNKTVSTLTOTENSRAECKTLQEL-----EKCAINEEDSKLEE 302  
QY 156 LKQKSSN-----SRKRKRSAGPTVPDRNDGIPDSLEVEGYTVDVQVNR 200  
DB 303 LKHNVANYSDAIVHKOKLIEDLSTRISEFDN-----LKSERDTLSIKNEKLEKLLRNTI 356  
QY 201 -TFUSPWISNIHEKGLTKYKSS-----PEKMWASDPYSDPEX-----VTGRID----- 244  
DB 357 GSKLDSRTSNSELEEMVELKESNRTIHSQLTDAESKLSFEQENKSLKGSIDVYQNLS 416  
QY 245 -----KNVS---PEARHPLVAA-----YPIVHVDM----- 266  
DB 417 SKRWVQVSSQLEAEARSLAHATGKLAETNSERDFQNKIKDFEKIEQDLRACLNSSN 476  
QY 267 ---ENIILSKNEDOSTQNTD---SETRTISKNTST-----RHTSEVHGNAEVHASFDP 315  
DB 477 ELKEKSAIDKQDQELNLRQIKQEKQKVSSTQSSLSQSLQDRIDILNEKKKHEVYESQLNE 536  
QY 316 IGGSVSAGFSNSG-----STVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTG 370  
DB 537 LKGLQTEINSEHLSQSLTLAEKAEAVATNNELSESK---NSLQTL-CNA----- 585  
QY 371 TAPIYNVLPSTSLVGNQX-----TLATKAKENQSLQILAPNNYPSKNLAPALMAQKD 426  
DB 586 ---FOEKLAQSMQKENEQNFSSLDTSFKKLNEHQELENHQTITKQL-----KD 634  
QY 427 FSSPTITVYNNQFLELEKTKOLRLDTQVYGNATYVNGRVRVDVGSNWSVLPOIQE 486  
DB 635 TSS-----KIQQLQERANFEQKSTSLDENNDLTK-----LKLEE 672  
QY 487 TTARIIFNGKOLNIVERRIAAVNPSDPLETTKPM-TLKEALKI-----AFGFEPNQNLQ 541  
DB 673 SNKSLIKQEDVDVSLKNI-----QTLKEDLRKSEALRPSKLEAKNLREVIDNL- 722

QY 542 YQCKDIT-BFDNFDOQTSONIKQNLAEALNATNIYTVLDKIKLNAGNNILIRDKRPHYDR 600  
DB 723 -KGKHTLEAQRNDLHSSLSDAKNTWAILSSSELTXSSEDVKRLTANVETLTQDSK----- 776  
QY 601 NNIAVGADESVMKEAREVINSTEGLLNIDKDIR-----KILSGYIVEIE-----DT 649  
DB 777 -----AMKQSTSLVNSYQS---ISNLVHELDRDDHVMQSQNNLTLESSEKLTQDC 824  
QY 650 EGLKE---VINDRYDMLNISSLRQDKGKTFIDFKKYNDKLPYVINSNPKYKVVAVTKENT 706  
DB 825 ENITQNMWILDVQKLMKHVNVQESKV-SELKEVNGKLSLDLKNLLESSLNV-AISNDQ 882  
QY 707 IINP-----SENGDT-----STNGIKKI 724  
DB 883 ILTQLAELSKNYDSLEQESAQLNSGLKSL 911

## RESULT 8

WAPA\_BACSU ID WAPA\_BACSU STANDARD; PRT; 2334 AA.  
AC Q07833;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Wall-associated protein precursor.  
GN WAPA OR N17G OR BSU59230.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=93302506; PubMed=9316082;  
RA Foster S.J.;  
RT "Molecular analysis of three major wall-associated proteins of a  
RT Bacillus subtilis 168: evidence for processing of the product of a  
RT gene encoding a 258 kDa precursor two-domain ligand-binding  
RT protein."  
RL Mol. Microbiol. 8:299-310(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / BGSC1A1;  
RX MEDLINE=95219088; PubMed=7704263;  
RA Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;  
RT "Cloning and sequencing of a 29 kb region of the Bacillus subtilis  
RT genome containing the hut and wapa loci."  
RL Microbiology 141:337-343(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / BGSC1A1;  
RX MEDLINE=97124196; PubMed=8969509;  
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,  
RA Miwa Y., Fujita Y.;  
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome  
RT containing the lic and cgl loci, and creation of a 177 kb contig  
RT covering the gnt-sacXY region."  
RL Microbiology 142:3113-3123(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Frietz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.I., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Guiseppi G., Guy B.J., Haga K., Haeck J., Harwood C.K., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigovic C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Fumelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler E., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*.";  
RL Nature 390:249-256(1997).  
CC -!- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,  
CC MOTILITY, SECRETION OR DIFFERENTIATION.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED  
CC INTO THE MEDIUM.  
CC -!- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE  
CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE  
CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED  
CC MOTIF REPEATED 31 TIMES.  
CC -!- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME  
CC SIMILARITY TO THE REPEAT IN E. COLI Rhs GROUP OF PROTEINS (RhsA-D).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L05634; AAA22883.1; -;  
DR EMBL; D31856; BAA06656.1; -;  
DR EMBL; D29985; BAA06260.1; -;  
DR EMBL; D83026; BAA11683.1; -;  
DR EMBL; Z99124; CAB15959.1; -;  
DR PIR; S32920; S32920; wapa.  
DR Subtilist; BG10797; wapa.  
DR InterPro; IPR003305; CBM\_CenC.  
DR Pfam; PF02018; CBM 4\_9; 1.  
DR Pfam; PF05593; Rhs repeat; 14.  
DR TIGRfams; TIGR01643; YD repeat 2x; 17.  
DR Cell wall; Repeat; Signal; Complete proteome.  
KW SIGNAL  
FT CHAIN 1 228  
FT DOMAIN 29 2334 WALL-ASSOCIATED PROTEIN.  
FT REPEAT 504 869 3 X 101 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 504 605 1-1.  
FT REPEAT 636 736 1-2.  
FT REPEAT 769 869 1-3.  
FT DOMAIN 1021 2139 31 X 21 AA APPROXIMATE TANDEM REPEATS OF  
FT X(4)-G-X(4)-[YF]-X-D-X(2)-G-X(4).  
FT REPEAT 1021 1040 2-1.  
FT REPEAT 1042 1061 2-2.  
FT REPEAT 1063 1082 2-3.  
FT REPEAT 1083 1102 2-4.  
FT REPEAT 1109 1128 2-5.  
FT REPEAT 1129 1148 2-6.  
FT REPEAT 1150 1169 2-7.  
FT REPEAT 1174 1193 2-8.  
FT REPEAT 1199 1218 2-9.  
FT REPEAT 1219 1238 2-10.  
FT REPEAT 1646 1665 2-11.  
FT REPEAT 1667 1686 2-12.  
FT REPEAT 1690 1709 2-13.  
FT REPEAT 1711 1730 2-14.

FT REPEAT 1732 1751 2-15.  
FT REPEAT 1753 1772 2-16.  
FT REPEAT 1795 1814 2-17.  
FT REPEAT 1820 1839 2-18.  
FT REPEAT 1840 1859 2-19.  
FT REPEAT 1861 1880 2-20.  
FT REPEAT 1887 1906 2-21.  
FT REPEAT 1908 1927 2-22.  
FT REPEAT 1929 1948 2-23.  
FT REPEAT 1969 1982 2-24 (APPROXIMATE).  
FT REPEAT 1983 2002 2-25.  
FT REPEAT 2008 2027 2-26.  
FT REPEAT 2028 2047 2-27.  
FT REPEAT 2051 2070 2-28.  
FT REPEAT 2071 2090 2-29.  
FT REPEAT 2093 2112 2-30.  
FT REPEAT 2120 2139 2-31.  
SQ SEQUENCE 2334 AA; 258329 MW; B75138CCD278AA3 CRC64;  
Query Match 4.4%; Score 166.5; DB 1; Length 2334;  
Best Local Similarity 18.2%; Pred. No. 1.7;  
Matches 173; Conservative 126; Mismatches 336; Indels 317; Gaps 42;  
Qy 1 EVQENRLNSESSESSQGLLYGYFSLNFOA-----PMVVTSSSTGDLISPSSELENIP 54  
Db 269 EVERSDEVSKLEKNEEGYLLHLTADENMLKOPERYVPSIDPSTSLSVSDTFVMSAYP 328  
Qy 55 SENQYFQSAIWSGFIK--VKKSDEYFATSADNHVTM----- 89  
Db 329 ITTYSASSQKQDANLKAAYLVKTGYD-KITGTYAFMKFNLLKPIQNMVTVKATLKYVA 387  
Qy 90 -----WYDDEVINK-----ASNSNKIRLEKGR--LVOIKIQYOR 122  
Db 388 HSYGTGKATGLMD--TVNSNYDNAKVNTWTKPASKNIGKADVHKQGWASVDVTAAVKS 444  
Qy 123 ENTEKGLDFLKTWTDSONK---KEYISSDNLQLPKQKSSNRKSRKSTAGPTVP--- 176  
Db 445 WNSGGANYGFKLH-TWNGKEYWKLLS-----SANSANKPIEVTYTIPIKGN 491  
Qy 177 -----DRNDGIPD-----SLE-VEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSP 223  
Db 492 TPTIKAYHNGDSTGYFDISWKKVEGAKYKVIYNGKEYQAIASAGNV-----TSMSTKG 545  
Qy 224 EK-WSTASDPYDFEKTGTGRID-KNVSPEARHPLVAAYPIVHVDMENIILSKNE----- 275  
Db 546 KKIWPTSAETASKRYKL--HLDGKGAEALD-----SPYKNSGGSYATSKNYWIGVSA 599  
Qy 276 --DQSTQNTDSETRTISKNTSTRTSTSEVHGNAEYHASFDDIGGSVSAG-----FS 325  
Db 600 IFDQEGAMSAPAKPVPINVGKQAPSAKGYNNGNA-TGVFDLSWKAHSVATGYKYQOVFN 658  
Qy 326 NNSSTVAIDH--SLSLAGERTWAETWGLNTADTA-----RLNANIEYVNTGTA 372  
Db 659 KGKFETLDLGNQTSWTTKGKKIWIPTSAETKAGKYALHLKDGSGAELPINPGPTYKNAGGD 718  
Qy 373 PIYNVLPTTSLVLGKQNTLATIKAKENQLSQILAPNN-----YYPK-----NLA--- 417  
Db 719 GAKENYSFKIIAYNKDGEATASPAAPALPDIAAPKPNVTGYLYTNTKSSQTGVNLIWEK 778  
Qy 418 -----PIALNAQKDFSTPI-----TMNVTNQFLELEK-----QLRL 450  
Db 779 VQNAKGYKVIYNGKEYQSFVDGADHWTQNKNIWPTSEETKAGSYKLHLDGKGELAL 838  
Qy 451 DTQVYGN-----IATYN--PENGVRVVDVTSNWSSEV-L 481  
Db 839 DPSPVNNANGNYKGVKQVYFVLVAYDANGETIPTAPFNPTFHEGAEFLGTEEVNLIIDI 898  
Qy 482 P--QIETTARIIFNGKDLN-----VERIAAVNPSPDI-----ETT----- 517  
Db 899 PSGLNGATGNVIVNEEDLSIDQRGFLGLSRTYNSLDSSDLHFGQGWYDAETVISTD 958  
Qy 518 -----KPDWTLKEALKIAPG-FNEPENGNIQYQKDIETDFDFNFDQDTQSNIKNQLAEL- 569

Db 959 GGAMYIDEDATTHRFKAKDGTVPPTGYVLELTETADQFILTVDQTNAYFNKGGKQLQ 1018  
Qy 570 -----NATNIYTLVDKIKLNARONILIRDKRFHYDRN-----NIAVGADES 611  
Db 1019 KVVVDGHNNATVYNDKQNLTAITDASGRKLTFTYDENGHVTSITGPKNKVTVSYENDL 1078  
Qy 612 VKE-----AAR--EVINSSTEGILL 629  
Db 1079 LKKVTDGTVTSYDVSSEGLVQVQVSAKTEAKPVFTFYQVSGHRAKAINAKKETVY 1138  
Qy 630 NIDKDI-----RKILSGY-----IVEIEDTEGLKEVINDRYDMLNI 665  
Db 1139 SYDADKKTLLMTPNGRKVQVGYNEAGNPIQVDDAEGLKITNTKYEGNNV 1190  
RESULT 9  
ID YJL3\_YEAST STANDARD; PRT; 1803 AA.  
AC P47024; P87192;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-MAY-2000 (Rel. 39, Last annotation update)  
DE Transposon Ty4 207.7 kDa hypothetical protein.  
GN TY4B OR YJL113W OR J0780.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / FY1679;  
RX MEDLINE=97103775; PubMed=8948101;  
RA Czepluch C., Kordes E., Pujol A., Jauniaux J.-C.;  
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X  
reveals 19 open reading frames including URA2 (5' end), TRK1, PPS2,  
RT SPT10, GCD14, RPE1, PHO85, NCA3, ASF1, COT7, GZF3, two tRNA genes,  
RT three remnant delta elements and a Ty4 transposon.";  
RL Yeast 12:1471-1474 (1996).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z49389; CAA89409.1; -;  
CC SGD; S0003649; YJL113W.  
CC InterPro; IPR001584; Rve.  
CC InterPro; IPR001878; Znf\_CCHC.  
CC Pfam; PF00685; rve; 1.  
CC SMART; SM00343; Znf\_C2HC; 1.  
CC Transposable element; Hypothetical protein.  
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;  
Query Match 4.3%; Score 164; DB 1; Length 1803;  
Best local similarity 19.5%; Pred.No.1.6; Matches 270; Gaps 41;  
Matches 162; Conservative 123; Mismatches 274; Indels 270;  
Qy 1 EVKQENRLNESSSQGLGYFSLNFAQPMVVTSTGDLSPSSSELENIPSENQYF 60  
Db 1078 QLKAKTN---HETSPFKESIG---TNVVFNTNNEISLTKGTSLPIKTESINHH--- 1128  
Qy 61 QSAIWGFIKVKSDSETATSDATSDNHTVMVDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 1129 -----SNDY-----STNKVE-----KF 1140  
Qy 121 QREN---PTEKGLDFKLYWTDSONKEVTSNDNLQPLKQKSSNRKRSTSGPTVP 176  
Db 1141 EKENHHPPPIEDIVDM-----SPQDWMESNCQDGNLKLKVTDKNVPTDNGTINVSPR-- 1193  
Qy 177 DRNDGIDPDSLEVEGYVDVKNKRTFLSPWISNHEKK-----GLTKY----- 219

Db 1194 -----LEQNEIASGSPVQTVKSAFLNKEFSLNKKRKHRRKHNKSNITSVELEDRKKR 1247  
Qy 220 -----KSSPEKSTASDPVSDPEKVTGRIDKNVSPEARHPLVAAPIVHVDMENILSK 273  
Db 1248 SKNRVKLIPDNNEVTSAPKIRAIYNEALSKNPDLKERHEYKQAY---HKELQNLKDWK 1304  
Qy 274 NEQSTQNTDSETRTISKN---TSTSTHTSEVHGNAEVHASPFDDIGGSVASGFSNSNST 331  
Db 1305 VFDVDVKYSRSE---IPDNLIVFTNTFTFKRNG---IYKARIVCRGDTQSPDTYSVIT 1358  
Qy 332 VAIDHS-----LSLAGERT-WAETMGLNTADT-ARLNANIRVYNTGTAPIYVLP--TTSL 383  
Db 1359 ESLNHNHIKIFLAIANNRNMFMKTLIDINHAFLAKLEE-----IYIPHPHRRRC 1408  
Qy 384 VLGNQTLATIKAKENQ-----LSQI-LAPNNTYP-----SKNLAFIALNAQKDFS 428  
Db 1409 VVKLNKALYGLKQSPKEMNDHLRQYLNGIGLKONSYPGLYQYTEDKXLM-IAVYDDCVI 1467  
Qy 429 STPTIMYNQFL-ELEKTKOLR-----LTDQVYGNATYFENGKV----- 469  
Db 1468 AASNEQRLDEFINKLSNFELKITGLIDVDLTD-ILGMDLVYKRLGTIDLTLSFIN 1526  
Qy 470 RVTGGSNWSVLPOIQTETARIIFNGKOLNVERRIAANVP-SDPLETTKPD-----MTL 523  
Db 1527 RMD--KKYNEELKKIRKSIPHMTYK-----IDPKKDLVQMSSEEFROGVKL 1573  
Qy 524 KEALKIAFGNPNGLNQYQGDITFDFNFDOOTSONIKNQLAELNATNIYTVLDKIKL 583  
Db 1574 QQLL-----GELNYV-RHKCRYDIEFAVKVVARLVNYPHERVFIYKIIQYL-- 1620  
Qy 584 NAKNHLIRDKRPHYDRN-----NIAVGADES VVKEAHRE-----VI-----NS 622  
Db 1621 -----VAYKDIGHYDRCDNKKKVAITDASVGEYDAQSRIGVILYGMNIFVYSNK 1675  
Qy 623 STEGLLANIDKIRKILSGV-----IVEIEDT---EGLK----- 653  
Db 1676 STNRCVSTAEALHAYIEGYADSETKVTUKELGEGDNDNDIVMTSKPAIQGLNRSYQ 1735  
Qy 654 -----EVINDRYDMLNISLRQDGKFTI-----DFPKY 681  
Db 1736 PKERFTWIKETIIKEIKESIKLLKTKGNIADELLTKPVSASDPKRF 1784  
RESULT 10  
ID P115\_MYCHR STANDARD; PRT; 979 AA.  
AC P41508;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE P115 protein.  
OS Mycoplasma hyorhinis.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2100;  
RN [1]  
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=91138990; PubMed=1825306; Wise K.S.;  
RA Notaricola S.M., McInosh M.A., Wise K.S.;  
RT "A Mycoplasma hyorhinis protein with sequence similarities to  
RT nucleotide-binding enzymes.";  
RL Gene 97:77-85 (1991).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- DOMAIN: Consists of two putative central coiled-coil regions  
CC flanked by putative globular regions at the N- and C-termini.  
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA  
CC P115.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M34956; AAA25423.1; ..

DR PIR; JQ0894; JQ0894; ABC transporter.

DR InterPro; IPR003433; GTP-binding\_dom.

DR InterPro; IPR005289; GTP-binding\_dom.

DR InterPro; IPR003405; SMC C.

DR InterPro; IPR003395; SMC N.

DR Pfam; PF02483; SMC C; 1.

DR Pfam; PF02463; SMC N; 1.

DR TIGRfam; TIGR00650; MG442; 2.

KW ATP-binding; Coiled coil.

FT NP\_BIND 32 39 ATP (POTENTIAL).

FT DOMAIN 169 224 COILED COIL (POTENTIAL).

FT DOMAIN 231 400 COILED COIL (POTENTIAL).

FT DOMAIN 569 821 COILED COIL (POTENTIAL).

FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).

FT SEQUENCE 979 AA; 110566 MW; 30D51C56B56280F4 CRC64;

Query Match 4.3%; Score 162.5; DB 1; Length 979;

Best Local Similarity 18.9%; Pred. No. 0.81;

Matches 175; Conservative 133; Mismatches 341; Indels 277; Gaps 40;

QY 13 ESSSQGLGYFSDLNFAQPMV-----TSSTTGDLSIPSSSELENIPSE 56

DB 51 EQSAKQLRLGMDVIFAGSKTVKPOEKAMVLTFKNEDAIETKQIPTISLLKRGQGT 110

QY 57 NOYFQS-----AIWGFHKVKSDEYTFATSNADNHVWVDDQE--VINKASN 102

DB 111 NEFYNDQPVRYKDKIKNLAVESG---ISKSLAISQGTISEIAETPQRKAVIEEAA 167

QY 103 SNIRLEKGRLYQIKYQRENPTKGLDFKLYWTDSONKKEVSSDNLQLPELQKSSN 162

DB 168 TSYKLDKEAOKLRL-----TNDAD-KLQGAKELEKQVNSLDK-QASKAKIYLEK 219

QY 163 SRKRRSTAGTPVDRD-----NDGIPDSLEVEGYTVDVK-NKRTFLSPWTSNIHEKG 215

DB 220 SKALESVEGLVINDLNFNFKLNLTSLLEVEQQRNDLELNLTQYESSISQTVH--- 275

QY 216 LTKYKSPKWSSTADPYSDFEKVTGRID--KNVSP-----EAPPLVAAVPIHV 264

DB 276 ---FKTEVE-----SQIETSKLNLKNALSEINLQEAIEERKLIISGEIV-V 322

QY 265 DMENII--LSKN-EDOSTQNTDSEITISKNTSTSTH----- 300

DB 323 DQKTKIEIKKQVESLKIQINASKQREIELDQQLRLNKAQNSKLQENDINKEIGVLE 382

QY 301 --SEVHGNAEVHASFDPDGGSVAGSFNSNSTVAIDHSLSLAGERTWAE---TWGLNTA 355

DB 383 KKSAAAANILKQPFENKSFSLSGIKTKIKNSFLDFGVIGLASELFKVESEFSLAETV 442

QY 356 DTAELN-----ANIRYVNTGTAPIYVNLPTTSL-----VLG 386

DB 443 LGAALNIVMKTSEDVQLQADFLKKNLSGKATIPITSIKEREVEDHLLVLKGQKFLG 502

QY 387 KNOTLATIKAKNQ-----LSQILAPNNYPSKNLAPI-----ALN 422

DB 503 VAKELIEFDQFNKFLGFLGLGNILVVDVNDVANRIAKILDPKHYTIVLSGLDFRPGGTT 562

QY 423 AQKDFSTPTIMYNQPLEKTKQLRLDQDQ-----VYGNIATVNPENGVRVD 472

DB 563 GGSKLERTSI-LNYDIKIK-EHTNWLKFAEDQIHLKIKQOITYNIETVNSTIQOVKIE 620

QY 473 TGS-----NWSEVLPOIQE--TTARIIFNGKDLNLVERRIAAV----- 508

DB 621 ANSINSKLNILNEELNKLNASIEFKQEQQEDQESLNSPSEKLN-IEKQISTLIELN 679

QY 509 NPSDPL-----ETTKEDMTLK-----EALKIATGFNEPNCNLOVQCK 545

DB 680 SKDRLTNLISEQKGETKKQELDAKRLKNTQHSDSITEQNAKFLVEQNKRLSEHK 739

QY 546 DITE-----FDNFQDQTS--QNIKNQALNATNIYVLDKILNKNMILIRDKRF 596

DB 740 LTLAASAEQYSLDLIDIEQARHFVDSLKKELKELGNVNLNLEAITEFEVNRQ---YOEKKQ 795

QY 597 HYDRNNIAGADESVVKEAHRVINSSTEGLLNIDKD-----IRKILSGYIIVEIDTEG 651

DB 796 YIELITAKSKIEBAISDLDKIINKTTE--IVNLVNNENMVQKFGGKAEIHFT-- 851

QY 652 LKEVINDRYMLNIS---SLRQGKTFIDPKND-----KLPIYI-- 699

DB 852 -----DKNDILNSGVEISAQPPKTKIKNLRLFSGGEKAIISLLFAILKARPIPLCIL 905

QY 690 -----SNPNVKNVYAVTKENT 706

DB 906 DEVEALDESNNVRYVEFLKLLKENT 931

RESULT 11

BXD\_CLOBO STANDARD; PRT; 1276 AA.

AC P1321; 1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, last sequence update)

DT 01-NOV-1990 (Rel. 16, last sequence update)

DT 28-FEB-2003 (Rel. 41, last annotation update)

DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)

DE (Bontoxilysin D).

GN BOTD.

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1491;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BYD/-3.

RA MEDLINE=91016853; PubMed=2216736;

RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,

RA Kozaki S., Krieglstein K., Henschen A., Gill D.M., Niemann H.;

RT "Nucleotide sequence of the gene encoding Clostridium botulinum

RT neurotoxin type D.";

RL Nucleic Acids Res. 18:5556-5556(1990).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CB16;

RX MEDLINE=93042276; PubMed=1420572;

RA Suragawa H., Ohyama T., Watanabe T., Inoue K.;

RT "The complete amino acid sequence of the Clostridium botulinum type D

RT neurotoxin, deduced by nucleotide sequence analysis of the encoding

RT phage d-16 phi genome.";

RL J. Vet. Med. Sci. 54:905-913(1992).

RN [3]

RP PARTIAL SEQUENCE.

RC STRAIN=D-SA, and D-1873;

RX MEDLINE=89339741; PubMed=2668193;

RA Morishita K., Syuto B., Kubo S., Oguma K.;

RT "Molecular diversity of neurotoxins from Clostridium botulinum type D

RT strains.";

RL Infect. Immun. 57:2886-2891(1989).

RN [4]

RP IDENTIFICATION OF SUBSTRATE.

RX MEDLINE=94230352; PubMed=8175689;

RA Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F.,

RA Roques B., Fykse E.M., Suedhof T.C., Jahn R., Niemann H.;

RT "Cleavage of members of the synaptobrevin/VAMP family by types D and

RT F botulinum neurotoxins and tetanus toxin.";

RL J. Biol. Chem. 269:12784-12772(1994).

CC -!- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER

CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED

CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD

CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT

CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC

CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-LEU-61 BOND OF

CC SYNAPTOBREVIN-1 AND -2.

CC -!- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the

CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No

CC detected action on small molecule substrates.  
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity). (L) and a  
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a  
CC heavy chain (H). The light chain has the pharmacological activity,  
CC while the N- and C-terminal of the heavy chain mediate channel  
CC formation and toxin binding, respectively.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of  
CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.  
CC -1- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D  
CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE  
CC BACTERIOPHAGE.  
CC -1- SIMILARITY: Belongs to peptidase family M27.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X54254; CAA38175.1; -;  
CC DR ENBL; S49407; AAB24244.1; -;  
CC DR PIR; S11455; S11455.  
CC DR HSP; P10845; 3BTA.  
CC DR MEROPS; M27.002; -;  
CC DR InterPro; IPR008985; ConA like lec\_g1.  
CC DR InterPro; IPR002160; Kunitz legume.  
CC DR InterPro; IPR006025; Pept\_M27\_2S.  
CC DR InterPro; IPR000395; peptidase\_M27.  
CC DR Pfam; PF01742; Peptidase\_M27; 1.  
CC DR PRINTS; PR00760; BONTOKILYSIN.  
CC DR ProDom; PD001963; Bontoxilysin; 1.  
CC DR PROSITE; PS00142; ZINC PROTEASE; 1.  
CC KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
FT CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.  
FT FT 443 1276 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.  
FT METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 230 230 BY SIMILARITY.  
FT FT 233 233 ZINC (CATALYTIC) (BY SIMILARITY).  
FT DISULFID 437 450 INTERCHAIN (PROBABLE).  
FT FT 15 16 ND -> PV (IN STRAIN D-1873).  
FT VARIANT 17 18 ND -> LQ (IN STRAIN D-1873).  
FT VARIANT 452 452 K -> Q (IN STRAIN D-SA).  
FT VARIANT 457 457 R -> T (IN STRAIN D-SA).  
FT VARIANT 457 457 R -> F (IN STRAIN D-1873).  
FT VARIANT 462 462 A -> D (IN STRAIN D-1873).  
FT VARIANT 489 489 K -> N (IN STRAIN CB16).  
FT VARIANT 644 644 N -> K (IN STRAIN CB16).  
FT VARIANT 1122 1122 Q -> R (IN STRAIN CB16).  
FT SEQUENCE 1276 AA; 146871 MW; C1E50F46C8233E2 CRC64;  
SQ  
Query Match 4.3%; Score 161; DB 1; Length 1276;  
Best Local Similarity 20.6%; Pred.No.1.4;  
Matches 164; Conservative 117; Mismatches 291; Indels 224; Gaps 41;  
QY 44 SIPSELENIENQYFQSAIWGFIKVKSDVETPATSDNHTVMVDDQEVINKASNS 103  
DB 310 TIPSSWTSNI---DKY--KKLFSEKYNFDKNTGNTGVNIDKNSLYSDLTVMSEVVYS 364  
QY 104 NKIRLEK-----GRLYQIKYQYQRENPTEKGLDFKLYWTSQNKKEV 145  
DB 365 SQYVKNRTHYFSRHYLPVPANILDDNIIYTIIRDGF--NLTKGFNIE----- 409  
QY 146 ISSNLQL-PELKQSSNSRKRSTAGTPVDRDNGIPDSLEVEGYTVDKNKRFLS 204  
DB 410 NSGQNIERNPALQKLSSVVDLTFKVCLRTKNSRD-----DSTCIKVNRL--- 458  
QY 205 PWISN-----IHEKGLTKYKSPKWKSTASDPYS-DPEKVTGRIDKQVSPKARHPLV 256  
DB 459 FVADKDSIQSEIFENKIITD-ETNVQNY---SDKFSILDESILDGQVP--INPEIVDPL- 511

QY 257 AAVPIVHVDMENILSKNE-----DOSTQNTD-----SETRTISKNTS--TSRTHTSEV 303  
DB 512 ----LPNVNNEPLNLPGBEIVFYDDITKYVDYLSYVYLESQKLSNNVENITLTSVEEA 567  
QY 304 HGAENVASPF-----DIGGSVAGSNSNSSTVAIDHSLSLAGERTWAETM----- 350  
DB 568 LGYSNKIYTFPLSLAEKVKVQAGLFLNWANEVVEDFTTNIMMKDITLDKISVSVIIPY 627  
QY 351 --GLNTADTA-RLNANIRVVTG-----TAPIYVNLPTTSLVLGKQNTLATIK 395  
DB 628 IGPALNIGSALRCNFCAATAGVAFLLBGFPEFIPALGVFTFYSSIQEREKIITIE 687  
QY 396 -----AKENQLSQILAPNNY-----YPSKNLAPIALNAQDFSSPITM 434  
DB 688 NCLEQRVKRWKDSYQWVSNWLSRITTFQFNHINYQYDS-----LSYQADAIAKIDL 740  
QY 435 NYNQFLEK-----TKOLRLDTDQVYGN-----ATYFENGVRVDTGSN 476  
DB 741 EYKISSGDKENIKSQQVENIKNSLDVKISEAMNINKFIRECSVTYLYFKN----- 790  
QY 477 WSEVLPQIQTETARIIFNGKD--LNL-----VERRIAAVNPSDPLETTKPDMTL 523  
DB 791 --MLPKVIDELNKFDLRTKTLEINLIDSHNIIIVGEVDRLKAKVNES--FENTMP--- 841  
QY 524 KEALKIAGF-PNEENGNOYQGGKIDTEFDNFDOQTS-----ONIKNOLAELNATNI-YTV 577  
DB 842 -----FNIFSYTNNSLL---KDIINEYFNSINDSKILSLQNKKALVDTSGYNAEVRV 891  
QY 578 LDKIKLNAKNNILIRKRFHYDRNNIAGVADSVKVEAHREVINSSTEGILLANDKDIRK 637  
DB 892 GDNVQLN---TIYTNDPKLSSGDKIIVLNNNIIYSA---IYENSVSFWIKISKDLTN 945  
QY 638 ILSGYIV--ETEDTEGKKEVINDRYDMNLSSLRQDGKTIDPKYNDKPLV---ISNP 692  
DB 946 SHNEYTTINSIEQSGWKLCIRNG---NIEWILQDVN-----RKYSKLIFDISELSHT 996  
QY 693 NYKVNVAVTKEHTII 708  
DB 997 GYTNKWFPTITNNIM 1012  
RESULT 12  
RBPI\_PLAVB  
ID RBPI\_PLAVB STANDARD; PRT; 2869 AA.  
AC Q00798;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Reticulocyte binding protein 1 precursor.  
GN RBPI.  
OS Plasmodium vivax (strain Belem).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=31273;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92315338; PubMed=1617731;  
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.,  
RT "A reticulocyte-binding protein complex of Plasmodium vivax  
merozoites".  
RL Cell 69:1213-1226 (1992).  
CC -1- FUNCTION: Involved in reticulocyte adhesion. Specifically binds to  
human reticulocyte cells.  
CC -1- SUBUNIT: Homodimer (potential).  
CC -1- SUBCELLULAR LOCATION: Membrane-bound.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC



DR	EMBL; M88097; AAA29743.1; -	
DR	Malaria; Receptor; Signal; Transmembrane.	
KW	SIGNAL	1
FF	CHAIN	17
FT	DOMAIN	18 2869
FT	TRANSMEM	2808 2826
FT	DOMAIN	2827 2869
FT	SITE	1030 1032
FT	SITE	2599 2601
SSQ	SEQUENCE	2869 AA; 330213 MW; B9DBE442205BCFF CRC64;
	Query Match	4.2%; Score 159.5; DB 1; Length 2869;
	Best Local Similarity	18.4%; Pred. No. 4.8;
	Matches 146; Conservative 142; Mismatches 297; Indels 209; Gaps 38;	
QY	10 NESSESSQGLLYYPSD-----LNQAPMVVTSTTCDLIPSELENIISENQYFOS 62	
QY	1944 NSMYSESWTFLANYFLTSDIAKISGVEFNKSNFKTDLELFIYSV-----ISNSNEL---- 1996	
QY	63 AIWSGFIKVKKDEVTFTFATSDAHVHTVMWDDQEVINKASNNKIRLEKGRLLQIKIYQYOR 122	
QY	1997 -----LKKIEQ-----DSNDVIQERESQLAKADADIYV-IKLN 2032	
QY	123 ENPTEKGLDFKLYWTDSONKKEVISSDNIQLPELKOKSNSRKRTSAGTPVDRDNG 182	
QY	2033 E-----FNEKLEAEKKEEVS-----EKVREALKRLS----- 2060	
QY	183 IPDSLEVEGYTVDVKNKRTFLSPWISNIHE---KKGLTKYKSSPKWSTASDPYSDPEK 238	
QY	2061 -----QVGIRCHFENFRLH-----DNTSELENLKNQVIYR---DKSERESGLQEMEN 2108	
QY	239 VTGRIDKQVSPARHPLVAAPIVHVDMENIL-----SKNEQSTQNTDSETRTISKWVS 294	
QY	2109 EMNTYNSNIT-----QLEGIVVSAGESKEDIKELRSNEEMENISEKIS 2152	
QY	295 TSRTHTSEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHSLSL-----AGERTWAEITM 350	
QY	2153 TDSKVIENSTID---ELYKGNQCA-----HWISLISYANMKTSEKLI 2196	
QY	351 GLNTADTARLNANIRYVNTGTAPIYVNLPTTSVLGNQNTLATIKAKENQLSQILAPNNY 410	
QY	2197 MIN-KEKENTEKCDVYIKONSSSTGDGYETLKGFGYSGKLTFFSS-----ASEIVQNADT 2248	
QY	411 YPSKNLA-----PIALNAQKQDPSSTPIITMYNQFLEL-----EKTQLRLTDOV 455	
QY	2249 Y-SVNFATKEKESLNARIDIKELYLFHQNSDISIVEGGVQNMLAYDKLNEKREMDL 2307	
QY	456 YGNIAITYNFENGVRVDTGSNWSEVLPIQIETTARIIFNGKDLNLVERRIAAYNPSDPLE 515	
QY	2308 YRNISETKLQMEHSTDVPKPMIELHKGMMNET-----NNKSLLLEKELKSVN---DHM 2359	
QY	516 TTPEDMTLKEALKI-----AFGFNEPNCLOQCKDITEPDFNF--DOOTSQNTKNQLAEL 569	
QY	2360 SMEAEM-IRKGLKYTPESVQNIINIVSIEAEVKTLEEIDRDDYGDNIQVIVEEHKQFSL 2418	
QY	570 -NATNIYTLVDKILAKAKMILIRDKEFHY---DRNNIAGVDESIVKEAHRVINSST- 624	
QY	2419 IDRTN--ALMDDIEIPKEN-----NYNLMEVNTETIHRVNDYIEKITNKLVOAKTE 2468	
QY	625 -EGLLNI---DKDIRKILSGYIVIEIDTEGLKE---VINDAYDM-----LNTSSLR 669	
QY	2469 YEOILENIKQNDMLQWIFIKKVSIIIEYFNVVKKKEISINDLYEQERLLKIGEHLDK 2528	
QY	670 QDGKTFIDFKYNDKPLPLYISN-----PNYKNVYAVYTKENTIINPS-----ENGDT 717	
QY	2529 RNVTETLSSYEIDQKEMMMSKULLEKKSQWNY-TS-YELEREANEINRDAKQIKDDTI 2587	
QY	718 TNGIKTILIFSKG 731	
QY	2588 LNSYLEAAI-QKRG 2600	

RESULT 13

```

S155_YEAST                                PRT; 1000 AA.
ID S155_YEAST                               STANDARD;
AC P43612;
AD
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S1T4-associated protein SAP155.
DE SAPI55 OR YFR040W.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
OX [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP MEDLINE=9620458; PubMed=8649382;
RX Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Arndt K.T.; a new family of proteins, associate and function positively
RT with the S1T4 phosphatase."
RL Mol. Cell. Biol. 16:2744-2755(1996).
RN [2]
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naïtou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Teuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae."
RC Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=96287652; PubMed=8696379;
RA Eki T., Naïtou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA Sasanuma M., Teuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of
RL chromosome VI from Saccharomyces cerevisiae.";
RY Yeast 12:149-157(1996).
CC -!- FUNCTION: Associates with the S1T4 phosphatase in a cell cycle
CC dependent manner. May be directly or indirectly involved in
CC S1T4-dependent functions in budding and in normal G1 cyclin
CC expression.
CC -!- PTM: Hyperphosphorylated in the absence of S1T4.
CC -!- SIMILARITY: Belongs to the SAPS family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U50560; AAC49303.1; -.
DR EMBL; D50617; BAA09279.1; -.
DR GeneBank; AF0194; -.
DR SGD; S0001936; SAPI55.
DR InterPro; IPR007587; SAPS.
DR Pfam; PF04499; SAPS; 1.
DR Phosphorylation; Ceil cycle.
KW CONFLICT 663 663
FT CONFLICT 674 689
N -> T (IN REF. 2 AND 3).
TYKSNMYRDECFQNN -> DLFKIKLYDTRIVSKIM (IN
REF. 2 AND 3).
DY -> EL (IN REF. 2 AND 3).
RVSSN -> PDIQVI (IN REF. 2 AND 3).
SO SEQUENCE 1000 AA, 114580 MW, AD4PFBFGC7CS588DS CRC64;

```

QY 1 EVKQENRLNESESSQGLGYFS--DL-----NFQAPMVVTSSTGD-----LSIPS 47

```
Db 90 EKEEERSNSESSTSTSSGSGTSTKTDLEEDISNATAPMMVTTKNLDNNGSIERMLVET 149
Qy 48 SELENIPSENQYFQSAINSGFIKVKSDYVATATSADNHVTMMVDD-OEVNKAANSNKI 106
Db 150 ELINELSRQNKLLDIFCGPFDKINKV-----NNMEYLVQLMCEKSKIKATV 203
Qy 107 RLEKGLYQIKIQYQRENPTKGLDFKLYWTSQNKVEISSDNIQLPELKQKSNRKK 166
Db 204 DLNN-----LIDYQBOQLD-----DSSQEDVVESDTEBEKEDNNNNKK 247
Qy 167 R-----STAGPTVPDRNDGIPDSLEVEGYVDVKNKTEFLSP-----NISNIHEKG 215
Db 248 RRRGSSFGNDNNDDDDANEDDESAYL-----TKATIISEIFSDIMWIS-----ES 299
Qy 216 LTKYSSPEK-WSTASDPYDFEKTGTRDKXVSPPEARPLVAAYPIVHVDMENIILSKN 274
Db 300 LVNQSYLNKWSIIINQ-----NFNSE-NSPLV-----PIFLKINQNLILLR- 341
Qy 275 EDQ-----STQWTDSET-----RTISKNT 293
Db 342 QDQYLMFIRTFERSFVDMKLVHVDISLLMDFKILSTOKIESPTGIIELVYQNLISKCL 401
Qy 294 STSRTHSEVHGAHVHASFDFDGGSVSAGFNSNSSTVAI-DHSLS--LAGERTWAETM 350
Db 402 SFLNNKESP-----ADIQACVGDFLKAIALSANAPLDDISIGPNSLTRQLASPEIAKLV 457
Qy 351 GLNTADTARLANIRVNTGTABIYVLPSTSLVGLKQNTLATIKAKENQLSQILAPNNY 410
Db 458 DI-----MINORGAALNT-----TVSIVI-----ELIRKNSDYDQVLLTTTITK 498
Qy 411 YPSKNLAPIALNAQKDFSSPTIMYVNOFLEKTKQLRLDQVYVGNATVNFENGRVR 470
Db 499 HPSNRDPVYGLLRKFSNHLSDFFQIILDIENDANIPLHENQLHEKPKLGFERFKV- 557
Qy 471 VDTGSNWEVLPOIQTETARIIFNGKDLNLVERRIAANVP--SDPLETTKPMTL--KEA 526
Db 558 -----VELIALLHCSNGLMNSKRAERARRDKVRSQLSHLQDALNDLSIEKEEQ 610
Qy 527 LKIAFG-----FNEPNCNLOYQGGDI-----TEFDNFD-----QOTSQNIKNOLAE 568
Db 611 LKTKHSTRTDHLKNNKINDNDNDDSDYGEIDESEFPIYINMKNIKLRTP 670
Qy 569 LNATNIYVLDKIKLNAMNILLRDKRFHYDRNRIAGVAD-----SSVKEAHREVIN-- 621
Db 671 TVGT-----YKSNYMIRD-----CPQNNELFTHPWNFWENFIDIIQOIFNGR 716
Qy 622 ---SSTGLILNID--KDIRKILSGYIVETE---DTEGLKEVINDRYDMLNISS---LR- 669
Db 717 MDFSNSFLVSLFNLSKSYQFTWIDVISDEKTDVSRFSPIRDPNFDKITTDFILRG 776
Qy 670 -QDGTTFIDFKKN-----DKLPLYISN-----PNYKVNYY 699
Db 777 YQDSYKFEYLRKNNLGYMGHIVLIAEYVVKFKLYKVDYISRYSSNLTQTEWQYSEVL 836
Qy 700 AVTK-----ENTIIINPSENGDTSN 719
Db 837 NETRMWYSKILGGSGYIDGNGNIIPOLPNTTVLTP--NGDASNN 880
```

## RESULT 14

```
FAT1 SCHPO
ID -FAT1 SCHPO STANDARD; PRT: 1385 AA.
AC O13735; QOUTJ0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
```

```
OX NCBI_TaxID=4896;
RN [1]
RC SEQUENCE FROM N.A.
PC MEDLINE=2184401; PubMed=11859360;
EX Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gellies A., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford J., Rutter S., Saunders R., Squares S., Stevens K.,
RA Skelton J., Simmons M., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Garzon A., Thode G.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Mueller-Auer S.,
RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler S., Gloux S., Lelaure V., Mottier S.,
RA Goffeau A., Cadieu E., Dreano S., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
RN [2]
RP IDENTIFICATION, AND GENE NAME.
EX MEDLINE=20143585; PubMed=10679021;
RA Jin H., Amberg D.C.,
RT "The secretory pathway mediates localization of the cell polarity
regulator Alp3/Budp3."
RL Mol. Biol. Cell 11:647-661(2000).
CC -!- FUNCTION: Involved in the organization and/or function of the
actin cytoskeleton.
CC -!- SIMILARITY: TO YEAST BUD6.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; Z97208; CAB10112.1;
DR EMBL; AL109770; CAB52420.1;
DR GeneDB SPombe; SPAC15A10.16;
DR GO; GO:0005884; C:actin filament; ISS.
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR InterPro; IPR005613; AIP3.
DR Pfam; PF03915; AIP3_1.
KW Coiled coil; Cytoskeleton.
FT DOMAIN 1009 1096
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40CC4FF7537 CRC64;
```

```
Query Match 4.2%; Score 159; DB 1; Length 1385;
Best Local Similarity 19.2%; Pred. No. 1.9;
Matches 147; Conservative 125; Mismatches 293; Indels 202; Gaps 33;
QY 26 DLNQAPMVVTSSTGDL-----SIPSSLENTIPSENQYFQSAINW-----GFIKVKKS 74
Db 326 DRKFSPLKTRTFSLTKSLDPTGTSLSXSPSLRKSFS--SFVQKDVYRSNLSRISQANRS 383
QY 75 DEYTFATSADNHVTMMVDDQDVINKAS-----NSN 104
Db 384 NVFPQATD-----VTRSVSDHRLSSINDGEVAPLPQRSTISSPNSPLSATVLPST 439
```

105 KIRLEKGRLYQIKIOYQREN-----PTEKGLDFKL-----YWTDSQNK 143  
440 PILLPRGRSSTLVNKKQFNADDGSLPNSIRTEYAAAPKLEDIADEVEDATISQR 499  
144 EVI-----SSNLQOLPELKQKSS---NSRKKRSTAGFTVPDRNDGI-----PDSL 187  
500 ELLERQIQABESSEDTSETISLQGLSLPQVSSTQOEIQPSSVPEAASNAIEAKEPAVT 559  
188 EVEGTVDVONKRTFLSPWISNIHEKKGITKYKSSPEKSTASDPYSFEKVTGRIDKNV 247  
560 ALESIT-----ERKEAPVISSEKIESGTSISTDTGGLANFENDSELEBLIQON- 613  
248 SPEARHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRITISKNTSTSRTHTEVHGNA 307  
614 NAEODEPSYKFKHYEYSSEE-----SGSEDFKSEKDTKGYVISNDDSTQVEEDSEKSTP 669  
308 EVHASFFDIGGSVAGFNSNSSTVAID-----HSLSLAGERTWAETMGLNTADTARLNAN 363  
670 NTGA-----SAKLINDPSSTITVSDVYKPKPASPEITEPPSSALVSATSPPTNVP 720  
364 I--RYVNTGTAPYINVLPTTSVLGKNTQTLATIKAKENQLSQILAPNNYPSKNLAPIAL 421  
721 IPVBAVHLSTA--FSTAPVSTIV-----SNISPLP-VAP--P 753  
422 NAKQDFSSPTIM--NYNQFLEKLT--KQLRLDQVYGNIAYNFENGVRVDTGSNW 477  
754 NVSGSPSETPKPEKVPVSGTEKALPKPLGVDTERY--FLRYNNQTRKVKVESPISN 811  
478 SEVLPOIQOETTARI-FNGK--DLNIVERRIAAVNPSDPLETTKPDMT-LKALKIAPGFN 534  
812 ANELGELFNVYKISFGSDSYELNIED-----PDKISVLLLEDLSLKYSLVSPWFK 864  
535 EPNGLQ--YOGKDIITFDFNFDQTSQNTKQNLAEALNANIVT-VLDKTKLNAKNMIL 590  
865 EQDANKKREDFHSGEVSATQHSQAQNTLDDHVTNTTHESPSSAFTAILERLKA----- 917  
591 IRDKRPHYDRNNIAVGADESUVKEAHREVINSSTGLLNLIDKQIRKILSGVIV-EIEDT 649  
918 -----IEQNIST-----NHTNDSAAKSESDHUKLANNFSPDSIDH 954  
650 EGLKEVINDRYDMLNISSLRQDKTFI-----DFKK-----YNDX 684  
955 KFYQVQKMNQ--LELASLKQISAAFFTRPLKIKDKFKKEINAFNEK 998

RESULT 15  
CAGA\_HELPJ  
ID CAGA\_HELPJ STANDARD; PRT; 1167 AA.  
AC Q9ZLT1;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Cytotoxicity associated immunodominant antigen (120 kDa protein)  
DE (CAG pathogenicity island protein 26).  
GN CAGA OR CAI OR CAG26 OR JHP0495.  
OS Helicobacter pylori J99 (Campylobacter pylori J99).  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Helicobacteraceae; Helicobacter.  
OX NCBI\_TaxID=85963;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99120557; PubMed=9923692;  
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
RA Trust T.J.;  
RT "Genomic sequence comparison of two unrelated isolates of the human  
RT gastric pathogen Helicobacter pylori."  
RL Nature 397:176-180(1999).  
CC -! FUNCTION: MAY BE NECESSARY FOR THE TRANSCRIPTION, FOLDING, EXPORT,  
CC OR FUNCTION OF THE CYTOTOXIN.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: AE001483; AAD06073.1; -  
DR PIR: B71924; B71924. Caga.  
DR InterPro: IPR005169; Caga.  
DR InterPro: IPR004355; IVSec\_caga.  
DR Pfam: PF03507; Caga; 1.  
DR PRINTS: PR01553; TYPE4SSCAGA.  
KW Antigen; Complete proteome. POLY-THR.  
FT DOMAIN 245 249  
FT DOMAIN 882 889 POLY-ASN.  
SQ SEQUENCE 1167 AA; 129729 MW; FDS8E6B81CEBD0F2 CRC64;  
Query Match 4.2%; Score 158.5; DB 1; Length 1167;  
Best Local Similarity 19.7%; Pred. No. 1.6;  
Matches 158; Conservative 149; Mismatches 308; Indels 187; Gaps 41;  
QY 3 KOENRLNLSSESSQGLL---GYFSDL-NF-----QAPMVVSTTGDLSIPSELE- 51  
DB 480 KADKALDREAKTTLOGSLKHGDMFVDFYDNFKYTNASKSPDKGVGATNG-----VSHLEA 534  
QY 52 -----NIPSENQYFQSAIWSGFIKVKKSDEYTFATSDAHNVTTMWDD-----QEVIN 98  
DB 535 GFESKVAVFENLNLNLAITSV---VRQLEDKLIAGKLSPOEANKLVKDFLSSNKELVG 590  
QY 99 KASNSNKIRLE---KRLYQIKIOYQRENPTKGLDFKLYWTDSDQNKKEVIVSSDNLQLEP 155  
DB 591 KALNFNKAFAEAKNTGNYDEVK---QAQKDEKSL-----KKERLEKDVAK--N 635  
QY 156 LKQSSNSRKRSTAGTVPDR-----DNDGIPDSLEVEGYTVDVKNKRTFLSPWISNI 210  
DB 636 LESKSGNKKVMEAKSQAQSKQDEIFALINKEANDARAI-AYAQNLLKIGIKRELSKLENI 694  
QY 211 HEKKGITKYKSSPEKSTASDPYSDEK-----VTGRI-DKNVSPSEARHPLVAAPYIV 262  
DB 695 N--KDLKDFSKSPDFENGKN--KDFSKAEETLKALKGSKVLDGLINPEWIS----- 741  
QY 263 HVDMENILSKNEDQSTQNTDSETRITISKNTSTSRTHTEVHGNAEVHASFFDIGGSVA 322  
DB 742 --KVENLNAALNEFKNGKDKSKVTOAK--SDLENSIKDVIINQKITDKVDNLNCAVSV 797  
QY 323 GFNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARLNANIRYVNTG-TAPIYVLPPT 381  
DB 798 AKATGDFSGV-----EQALADLKNFSKQLAQQAQKQEDNFTKNSALYQ----- 842  
QY 382 SLVLGKNTQTLATIKAKENQLSQILAPNNYPSKNLAPIALNAQKDFSSFTITWVYQPLE 441  
DB 843 SVKNGVNGTIV-----GNGLSKAEAT--LSKNFSDI-----KKELNKLGKFNNNNNG 890  
QY 442 LEKTKQLRLDQVYGNIAYNFENGVRVDTGSNMSEVLPTQIQTETARIIFNGKDLNLV 501  
DB 891 LENS-----TEPIYTOVA---KKVKAIDRLDQIASGLGDVGOQAASFLKHKDKVDDL 940  
QY 502 ERRIAAVNSDPLETTKPDV-----TLKEALKTAFGNEPENGNYOGKDIITEFDNPDQ 557  
DB 941 SKVGLSAN-HEPIYATIDDLGGPFLKRDHV-----DDLSKVLGSLREKQ 984  
QY 558 TSNQIKN---QLAELNAT---NIYTVLDKIKLNAKNILIRDKRPHYDRNNIAVGADES 611  
DB 985 LTQKIDNLNCAVSEAKASHFDNLQDMIDKLGKSTKQNV-----NLVYESAKKV 1033  
QY 612 VKSAHREVINSSTEG---LLNLI-DKDKIRKILSGYIVIEDTEGLKEVINDRYDMLNISS 667  
DB 1034 PLSLSAKLDNATNSHTRINSNVKNGTINEKATGMLTQ-KNSWLK-LVNDKIVAHNVGS 1091  
QY 668 LRQDKGTFIDF-----KYNID--KLPIYISNP-----NKNVNYAVTKENTIIIFS 711

Db 1092 APLSAYDKIGFNQKNMKDYSDSKFSTRLSNAVKDIKSGFYQFLTNIFSMGYSIMKASV 1151

Qy 712 ENGDTSTNGIKKILIFSKGYE 733

Db 1152 EHGKNTN-----TKGGFQ 1165

Search completed: May 3, 2004, 19:36:51  
Job time : 13.2539 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:17 ; Search time 34.6261 Seconds

(without alignments)  
6697.424 Million cell updates/sec

Title: US-09-848-909A-8  
 Perfect score: 3773  
 Sequence: 1 EVKQENRLNESESSSGLL.....TSTNGIKKILIFSKKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

```

maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database : SPTREMBL 25:\*

```

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvms:*
16: sp_bacteriap:*
17: sp_archeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3695	97.9	741	2	Q937W2	Q937W2 bacillus an
2	3695	97.9	743	2	Q937W3	Q937W3 bacillus an
3	883.5	23.4	876	2	Q32739	Q32739 clostridium
4	883.5	23.4	879	2	Q06498	Q06498 clostridium
5	881.5	23.4	876	2	Q9KH41	Q9KH41 clostridium
6	877	23.2	875	2	Q46221	Q46221 clostridium
7	819	21.7	721	2	Q86171	Q86171 clostridium
8	734	19.5	725	2	Q844J8	Q844J8 bacillus th
9	296.5	7.9	225	2	Q8KYK2	Q8KYK2 bacillus an
10	220.5	5.8	1881	16	Q8RGK2	Q8RGK2 fusobacteri
11	217.5	5.8	3468	5	Q31104	Q31104 plasmodium
12	205	5.4	4688	16	Q3PQ08	Q3PQ08 ureaplasma
13	202.5	5.4	1904	5	Q8IKS2	Q8IKS2 plasmodium
14	199.5	5.3	2849	5	Q8IHX4	Q8IHX4 plasmodium
15	194	5.1	1404	2	Q8RJN9	Q8RJN9 mycoplasma
16	191.5	5.1	675	16	Q897H6	Q897H6 clostridium

## RESULT 1

Q937W2	PRELIMINARY;	PRT;	741 AA.
ID	Q937W2		
AC	Q937W2; AC		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation)		
DE	Pag protein (Fragment).		
DE	PAG.		
GN	PAG.		
OS	Bacillus anthracis.		
OG	Plasmid pXOI.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae		
OX	NCBI_TaxID=1392;		
RN	[1]		
RP	SEQUENCE FROM N.A.		

RC STRAIN-Perrara;  
RA Adone R., Pasquali P., La Rosa G., Marianelli C., Muscillo M.,  
RA Fasanella A., Francia M., Ciuchini F.;  
RT "Sequence analysis of the genes encoding for the major virulence  
RT factors of bacillus Anthracis vaccine strain 'Carbosap'";  
RT submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

## ALIGNMENTS

17	191.5	5.1	1301	5	QWKS5	Q8wks5 plasmodium
17	191.5	5.1	2940	5	Q8HP9	Q8HP9 plasmodium
19	191	5.1	3248	5	Q8ICP9	Q8ICP9 plasmodium
20	189.5	5.0	2771	5	Q82616	Q82616 plasmodium
21	189	5.0	1936	5	Q8ISA6	Q8ISA6 plasmodium
21	189	5.0	2792	5	Q814R2	Q814R2 plasmodium
22	189	5.0	9904	16	Q8NW6	Q8NW6 staphylococ
23	188	5.0	1681	5	Q8IS48	Q8IS48 plasmodium
24	187.5	5.0	2747	5	Q8BJX9	Q8BJX9 plasmodium
25	187	5.0	2747	5	Q8BJX9	Q8BJX9 plasmodium
26	186.5	4.9	1301	5	Q8IHQ2	Q8IHQ2 plasmodium
27	185.5	4.9	2235	5	Q8IB09	Q8IB09 plasmodium
27	185	4.9	2269	5	Q82623	Q82623 plasmodium
28	185	4.9	2529	16	Q82579	Q82579 helicobacte
29	184.5	4.9	1859	5	Q8IC27	Q8IC27 plasmodium
30	184	4.9	3504	5	Q8IL45	Q8IL45 plasmodium
31	184	4.9	3504	5	Q8IL45	Q8IL45 plasmodium
32	183.5	4.9	1125	16	Q8PXY9	Q8PXY9 mycoplasma
32	183	4.9	3203	5	Q7WE7	Q7WE7 plasmodium
33	183	4.9	3203	5	Q7WE6	Q7WE6 plasmodium
33	183	4.9	3203	5	Q7WE6	Q7WE6 plasmodium
35	183	4.9	4152	2	Q9ZHL3	Q9ZHL3 haemophilus
36	183	4.9	4152	16	Q7VLS8	Q7VLS8 haemophilus
37	182.5	4.8	2444	16	Q81GX1	Q81GX1 bacillus ce
38	182	4.8	6713	16	Q9U54	Q9U54 staphylococ
39	182	4.8	6713	16	Q93186	Q93186 staphylococ
39	182	4.8	4095	2	Q7WZ13	Q7WZ13 haemophilus
40	181.5	4.8	2563	5	Q8THV4	Q8THV4 plasmodium
41	181	4.8	2647	5	Q9U4X0	Q9U4X0 plasmodium
42	181	4.8	10061	5	Q813Z1	Q813Z1 plasmodium
43	181	4.8	149	2	Q9RM77	Q9RM77 clostridium
44	180.5	4.8	1113	5	Q81500	Q81500 plasmodium
45	180	4.8	1113	5	Q81500	Q81500 plasmodium

Query Match 97.9%; Score 3695; DB 2; Length 741;  
Best Local Similarity 99.6%; Pred. No. 7.4e-168;  
Matches 719; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
1 EVKGNRLNESESSGCLGYFTSDLNFPQPMVVTSTTGDLSIPSSLENIPTSENQYF 60

Db 20 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTGDLSPSELENIPSENOYF 79  
QY 61 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 80 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 139  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLOPELKQKSSNSRKRSTASGTPVDRDN 180  
Db 140 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLOPELKQKSSNSRKRSTASGTPVDRDN 199  
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 240  
Db 200 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 259  
QY 241 GRIDKNSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSEPTRTSKNTSRTHT 300  
Db 260 GRIDKNSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSEPTRTSKNTSRTHT 319  
QY 301 SEVHGNAEVSASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 320 SEVHGNAEVSASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 379  
QY 361 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 380 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 439  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNGRVRVDTGNSWSEV 480  
Db 440 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNGRVRVDTGNSWSEV 499  
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPENGL 540  
Db 500 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPENGL 559  
QY 541 QYQKDIPTDFNFDQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRPHYDR 600  
Db 560 QYQKDIPTDFNFDQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRPHYDR 619  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRY 660  
Db 620 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRY 679  
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 720  
Db 680 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 739  
QY 721 IK 722  
Db 740 IK 741

DR EMBL: AJ413936; CAC93934.1; -.  
DR GO: GO:000576; C:extracellular; IEA.  
DR GO: GO:0046821; C:extrachromosomal DNA; IEA.  
DR GO: GO:0015070; F:toxin activity; IEA.  
DR GO: GO:0009405; P:pathogenesis; IEA.  
DR InterPro: IPR003896; Anthrax toxinB.  
DR Pfam: PF03495; Binary toxB; I.  
DR PRINTS: PR01391; BINARYTOXINB.  
KW Plasmid.  
FT NON\_TER 1 1  
FT NON\_TER 743 743  
SQ SEQUENCE 743 AA; 83363 MW; 292757AD2D5DESA6 CRC64;  
Query Match. 97.9%; Score 3695; DB 2; Length 743;  
Best Local Similarity 99.6%; Pred. No. 7.4e-168;  
Matches 719; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTGDLSPSELENIPSENOYF 60  
Db 22 EVKQENLLNESSSQGLGYFSDLNFOAPMVVTSSTGDLSPSELENIPSENOYF 81  
QY 61 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 82 QSAIWSGFIKVKKSDYTFATSNADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 141  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLOPELKQKSSNSRKRSTASGTPVDRDN 180  
Db 142 QRENPTKGLDFKLYWTDSONKKEVSSDNLQLOPELKQKSSNSRKRSTASGTPVDRDN 201  
QY 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 240  
Db 202 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPSDFEKT 261  
QY 241 GRIDKNSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSEPTRTSKNTSRTHT 300  
Db 262 GRIDKNSPEARPLVAAPYIVHVDMENIILSKNEDQSTQNTDSEPTRTSKNTSRTHT 321  
QY 301 SEVHGNAEVSASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 322 SEVHGNAEVSASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 381  
QY 361 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
Db 382 NANIRYNTGTAPIYVLPVTSVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 441  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNGRVRVDTGNSWSEV 480  
Db 442 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNGRVRVDTGNSWSEV 501  
QY 481 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPENGL 540  
Db 502 LPQIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNPENGL 561  
QY 541 QYQKDIPTDFNFDQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRPHYDR 600  
Db 562 QYQKDIPTDFNFDQTSQNIKNQLAELNATNIYTVLDKIKNAKWNILIRDKRPHYDR 621  
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRY 660  
Db 622 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRY 681  
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 720  
Db 682 DMLNISSLRQDGKTFIDFKYNDKPLIYISNPYKNNVAVTKENTIIINPSENGDTSTNG 741  
QY 721 IK 722  
Db 742 IK 743

RESULT 3  
O32739  
ID O32739 PRELIMINARY; PRT; 876 AA.

RESULT 3  
O32739  
ID O32739 PRELIMINARY; PRT; 876 AA.

Q937W3 PRELIMINARY; PRT; 743 AA.  
AC Q937W3;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE pag protein (Fragment).  
GN PAC  
OS Bacillus anthracis.  
OG Plasmid pXOI.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Carbosap.  
RA Adone R.; Pasquali P.; La Rosa G.; Marianelli C.; Muscillo M.,  
RA Fasanella A.; Francia M.; Ciuchini F.;  
RT "Sequence analysis of the genes encoding for the major virulence  
RT factors of bacillus anthracis vaccine strain 'Carbosap.'";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.



Db 321 ----DQKTVSRNTNSKTDANTAGVAINAYQNGF-----TGSITNTVSHHTENSTAVQNS 373  
QY 338 LSLAGERTWAETMGLNTADTARLANIRYVNTGTAPIYVNLPTTSLVGLKNGQTLTIKAK 397  
Db 374 ---NGE-SWNTSLINKGESAVINANRYVNTGTAPYKVPTTNLVL-DGDTLTITIKAQ 428  
QY 398 ENQLSCILAPNNYPSKNIAPALNAQDFSTPTITVNNQFLELEKTKQLRLDQDVYG 457  
Db 429 DNOIGNNLSPNTYPRKGLSPALNTMQFSLRIPINIDQKLDAGKQIKLETTQVSG 488  
QY 458 NIATYNFENGVRVDT-GSNMSEVLPIQIETTARIIPN-GKDLNLVERRIAAVNPSPDLE 515  
Db 489 N---YGIRKNSQOQIITEGNSWDYISQISLSASIILDTGSD--VFERRVTAKDSSNPED 543  
QY 516 TTKPDMTLKALKIAFGNEPNNGNLQYQKQITE--PDFNEDQQTSONIKNQLAELNATN 573  
Db 544 KT-PVLTIGEALEKAFGATKNGEILYFNGMPIDESCVELIFDGNANLIKRLNALNDKK 602  
QY 574 IYTVLDKIKLNAMNILLRD-----IAVGADESVMVKEAHREV----- 593  
Db 603 IYVW---QLRGMKILIKTSYFNNFDGYNFPSSWNSVDNODGLQNAANKLSGETK 658  
QY 594 -----KRFHY----- 598  
Db 659 IIVPMKLNPKRYVPSFGYLKNSSTNPITVNIKAKQKTYNLVSENDYKFSYEFETIG 718  
QY 599 -DRNN-----IAVGADESVMVKEAHREV-----I 620  
Db 719 RASNEIETLTSSGTFIDNLNLSITELNSTPEILKEPDIKVPSDQELI-DAHKYVADLSF 777  
QY 621 NSTEGELLN-----IDKDIRKILSGYIVIEB-DTEGLKEVINDRYDMLNLSRLQD 671  
Db 778 NOSTANYLDGLYFEPTQNKVELDYIQKYKVEATLEYSGRDKDGTCKELERNVTGDSNQ 837  
QY 672 GKT-FIDEKPKYNDKPLYSINPNY---KVVYVAVTKENTIIINPSENGDTSNGFKTILI 726  
Db 838 PKTYNVNFRSY-----FTSGENVMPYKLRIVATPEN-----KELLV 875  
QY 727 FS 728  
Db 876 LS 877

RESULT 5  
Q9KH41 ID Q9KH41 PRELIMINARY; PRT; 876 AA.  
AC 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
GN CdB.  
OS Clostridium difficile.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1496;  
RN [1] SEQUENCE FROM N.A.  
RP STRAIN=CCUG 20309;  
RA Chang S.Y., Song K.P.;  
RT "ADP-ribosylating Binary Toxin Genes of Clostridium difficile strain  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF271719; AAF81761.1; -.  
DR HSSP; P13423; IACC.  
DR GO; GO:0005576; C:extracellular; IEA.  
DR GO; GO:0015070; F:toxin activity; IEA.  
DR GO; GO:0009405; P:pathogenesis; IEA.  
DR InterPro; IPR003896; Anthrax toxinB.  
DR Pfam; PF03495; Binary toxinB; I.  
DR PRINTS; PR01391; BINARYTOXINB.  
SQ SEQUENCE 876 AA; 98793 MW; 366D62F352E745A5 CRC64;

Query Match 23.4%; Score 881.5; DB 2; Length 876;  
Best Local Similarity 31.6%; Pred. No. 5.5e-34;  
Matches 251; Conservative 140; Mismatches 279; Indels 125; Gaps 30;  
QY 4 QENRLNSESSESQGLLYFSDLNFOAPMVVTSSTGDLSPSSSELEN-IPSENQYFOS 62  
Db 40 KKKEIVNEDILPNGLMGYYFTDEHFQDLKMAPIDGNLKEPBEKKVLDKDKSDVKS 99  
QY 63 AIVSGIYKVKSDVTFATSAADNHYTMVDDQEVINKASNSNKKIRLEKGLYQIKIYQR 122  
Db 100 IRWTGRIILSKOGGYTLSTRDD-VLMQVNTSTI---SNTLKVNMKKKEKVKVLELOD 155  
QY 123 EN--PTEKGLDFKLYWTDSONKEVSSDNQLPELKQKSSNRKXKSTASQPTVP---- 176  
Db 156 KNLGSDNLNLSPLNY-ELDGMKKIPEENLFLRDY-----SNIKO-----DPFIPNNNF 205  
QY 177 -----DADNDGIPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSS 222  
Db 206 FDPKMSDWEDDLDTNDNIPDSYERNGYTI-----KOLIAVKWEDSFAE-QYKKTYSN 260  
QY 223 PEKWTASDPYDFEKVGTGRIDKIVSPEARHPLVAAYPVHVDMENIILSKNEDOSTONT 282  
Db 261 YLESNTAGDPYTYEKASGSFQKAIKTEARDPLVAAYPVGVGMKLIISTNEHAST--- 317  
QY 283 DSSTRISKNTSTSRTHTEVHGNAEVHASPFDIGGSVAGFSNSNSTVAIDHS----- 337  
Db 318 -DQKTVSRATNTSKTESNTA-----GVSVMVGYQNGFTANVTNYSHHTDN 363  
QY 338 ---LSLAGERTWAETMGLNTADTARLANIRYVNTGTAPIYVNLPTTSLVGLKNGQTLTI 394  
Db 364 STAVQNSGESWNTGISINKGESAYINANRYVNTGTAPYKVPTTNLVL-DGDTLSTI 422  
QY 395 KAKENQLSQILAPNNYPSKNIAPALNAQDFSTPTITVNNQFLELEKTKQLRLDQ 454  
Db 423 KAQENQIGNLSPGDTYPKKGLSPLALNTMQDQFSSRLIPINYDQLKLDAGKQIKLETTQ 482  
QY 455 VYGNIAATYPENGRVVDGTGSNWSVLPQIQETARIIFNGKDLNLVERRIAAVNPSPDPL 514  
Db 483 VSGNFGTKN-SSQGI-VTEGNSWSDVISQISASIIIDTEN-ESYERRVTAKNLQDPE 539  
QY 515 ETTKPDMTLKEALKIAPGNEPNNGNLQYQKQITE--PDFNEDQQTSONIKNQLAELNAT 572  
Db 540 DKT-PELTIGEALEKAFGATKKGGLLYFNDIPIDESCVELIFDDNTANKIKDLSKTLSDK 598  
QY 573 NIYTVLDKIKLNAMNILLRDKEFHY---DRNNIAVGADESVMVKEAHREVINSSTGL-- 627  
Db 599 KIYNV-----KLERGMNLIKTPITYFTNFDYNNYP--STWSNVNTTKDGLQGSANKLNG 652  
QY 628 -----LLMIDKDIRKILSGY-----IVEI-----EDTEGL-----KEVI 656  
Db 653 ETKIKIPMSSELKPKYKRVFSGYSKDPILTSNIIIVKIKAKEEKTDLVPEQGYTKFSYEPE 712  
QY 657 NDRYDMLNIS-SLRQDKTFIDFKYND--KLPLYISNPNYK-----VVVAVTKENT 706  
Db 713 TTEKSSNIEITLIGSGTIVLDNLSITELNSTPEILDEPEVKIPTDQEIIDAKHIIFADL 772  
QY 707 IINPSENGDTSNGI 721  
Db 773 NFNPS-TGNTYINGM 786

RESULT 6  
Q46221 ID Q46221 PRELIMINARY; PRT; 875 AA.  
AC Q46221;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
OS Iota toxin component Ib precursor.  
OC Clostridium perfringens.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1502;







Db 1 MESLGINNIYALDRIKNAKNNILVRDP-YHYDNGNIVGVDDSYLKNAKQILNWSDD 59  
Qy 626 GLLNIDKIRKILSGYIVIEDTE-----GLKEVINDYDMLNLISSLRQDG 672  
Db 60 GVSNLDEVDNALSGYMLQIKFSPNHLTNSPVITITLAGKDSGVGELYRVLS-----DG 113  
Qy 673 KTFIDFKYNDKPLIYINPNKVNVAATKRN-TIINPSENGDTSTNGIKIILFSGKG 731  
Db 114 TGFIDFKFENWRSVLV-DPGDDVYVAVTKEDFNVAITRDENGIA-NKLKNTLVLSGKI 171  
Qy 732 YEI 734  
Db 172 KEI 174  
RESULT 10  
Q8RGK2 PRELIMINARY; PRT; 1881 AA.  
AC Q8RGK2;  
DT 01-JUN-2002 (TremBLrel. 21, Created)  
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Hemolysin.  
GN FN0291.  
OS Fusobacterium nucleatum (subsp. nucleatum).  
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
OC Fusobacterium.  
OX NCBI\_TaxID=76856;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ARCC 25586;  
RX MEDLINE=2186694; PubMed=11889109;  
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
RA Bhattacharyya A., Battman A., Gardner W., Grechkin G., Zhu L.,  
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,  
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,  
RA Fonstein M., Kyripides N., Overbeek R.,  
RA "Genome sequence and analysis of the oral bacterium Fusobacterium  
RT nucleatum strain ARCC 25586.";  
RL J. Bacteriol. 184:2005-2018(2002).  
DR EMBL: AE010541; AAL94497.1;  
DR GO: GO:0008233; P:peptidase activity; IEA.  
DR GO: GO:0016020; C:membrane; IEA.  
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro: IPR008638; Haemagg.act.  
DR InterPro: IPR000508; Peptidase\_S26.  
DR InterPro: IPR001452; SH3.  
DR Pfam: PF05860; Haemagg.act; 1.  
DR Pfam: PF00018; SH3; 1.  
DR PROSITE: PS00761; SPASE\_1\_3; 1.  
KW Complete proteome.  
SQ SEQUENCE 1881 AA; 204375 MW; D635156A4EFA4877 CRC64;  
Query Match 5.8%; Score 220.5; DB 16; Length 1881;  
Best Local Similarity 20.8%; Pred. No. 0.04;  
Matches 175; Conservative 131; Mismatches 303; Indels 233; Gaps 37;  
Qy 30 QAPNWTST-----STTG-----DLSPSSSE---NIPSENQY-FQSAIWSGFI 69  
Db 246 QAPNATKGVGVSSKGVYKDTQAKRDKISSTETETIGSKLLAENAIKSGKTSNG 305  
Qy 70 KVKSDVETATSDADNHTVMVDQEVIN-----KASNSKIRLEKGR 112  
Db 306 QIRANNITINGVDSNLIPTNKDITISGNLKNVSSNINVKIEIENSKVWVEE-K 364  
Qy 113 LYQKI-----QYQENPTKGLDFKLYWT--DSQNKVEISSDNLQLPKOKSS 161  
Db 365 LSSTKITNLGNLSAKEIEKTIKFNISGKLFKSNITAKDFKNGEV-SSENITTTNLENSK 423  
Qy 162 -----NSRKRKSTTSAGTVPDRNDG----- 182  
Db 424 INVKENINNSIVNKTNAEITISKNLNTNLDNRGNIITINNVSSTGVANNKGLVGNNTIN 483

Qy 183 ---IPDSLEVEGYVDVKNKRTFFSPWISNTHKEKGLTKYKSSPEKWTASDPYDPKV 239  
Db 484 SQNLNTATVQKTLIDKKNKNSGKILSDNLTLDIFSSGNISAKVITTOELINGSII 543  
Qy 240 TGRIDKNSPEARHPVAAYPIVHVDMEN-----IILSKNEDQSTQNTDSTRITSKNT 293  
Db 544 SNNLSNNINNNKNIIFVGNLKIENNLSNVGIEGLNLSIENTGNTIKNKLTSQNL 603  
Qy 294 STSRTHTEVHGNAEVAHSAFFDGGVSAGFSNSSTVA---IDHSLSLAGERTWAETM 350  
Db 604 NNKK-----NTANVAGFLDVHNKISS-VGNIKAITMKTNNLDSGNIL-----TN 648  
Qy 351 GLNTAD-----TARLNANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKAKENQLSQ 403  
Db 649 SLTTAENINKSITAKNISQNLVNSGS-----VISDNIITVA-----DN 687  
Qy 404 ILAPNNYPSKNAPIALNAQKDFSTPITMNYNOFLELEKTKQLRLDQDVYGNATYN 463  
Db 688 IITNNIF-----ANEKISADKIS-NSNK-LVAKNTEITKLTND---GNIWVK- 730  
Qy 464 FENGVRVDTGSNWSEV-----LPQIQTARIIFNGKDLNLVERIAAVNPSPDPLE 515  
Db 731 -ENLKAKDITNSNTIKVGNLNTDKLQNSKTLIAKNINEKSLNNGINKITSLNANINTS 789  
Qy 516 TTRPDMTLKEALKIAFGFNEPNGNLQYQKDKITEFDFNFQDQTSQNKQLAELNATNY 575  
Db 790 DIKNNNGIIQAIK-----NINIKTSNDLKLDGKYTANDSLNINAKSLENNGN--- 836  
Qy 576 TVLD---KIKLNAKMILIRDKFHDNRNIAV-----GADESIVYKEAHRVINS--T 624  
Db 837 --LENDGKIKFNLTCNLNNKISSSNLNTANEISNNVNSIIGSEANLITANSKLN 894  
Qy 625 EGUILL-----NI-DKDIRKILSGYIVEIETEGLEKEVINDYDMLNLISSLRQ 670  
Db 895 EGNLLPFGIENKLTGNTTNGVTSLLGKLEAKDVNDKHIISDNDLTIDVNSITN 954  
Qy 671 DGKTF-----IDFKK--YNDKLPYISNPYKVNVAATKNTIINPSENGDTSTNGIK 722  
Db 955 KGLLYSTNNMKVDFKFNFLDKAEIYSSG-----DITIN-SENG-TFTNRVG 999  
Qy 723 KI 724  
Db 1000 DI 1001  
RESULT 11  
Q8II04  
ID Q8II04 PRELIMINARY; PRT; 3468 AA.  
AC Q8II04;  
DT 01-MAR-2003 (TremBLrel. 23, Created)  
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)  
DE Hypothetical protein.  
GN Pfli\_0371.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=3D7;  
RX MEDLINE=22255705; PubMed=12368864;  
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,  
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,  
RA Bisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,  
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,  
RA Perlea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,  
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,  
RA McPadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,  
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,  
RA Fraser C.M., Barrell B.,  
RA "Genome sequence of the human malaria parasite Plasmodium  
RT falciparum.";  
RL Nature 419:498-511(2002).

DR EMBL; AB014841; AAN35955.1; --  
KW Hypothetical protein.  
SQ SEQUENCE 3468 AA; 411913 MW; 506F7D62999BA7B1 CRC64;  
  
Query Match 5.8%; Score 217.5; DB 5; Length 3468;  
Best Local Similarity 19.0%; Pred. No. 0.12;  
Matches 159; Conservative 131; Mismatches 274; Indels 271; Gaps 37;  
  
QY 55 SENQYFQSAIWSGFIKVKKSDEYVTSADNHTVMWDDQEV---INKASNSKIRLESG 111  
DB 160 SENEY-----TKDDKYNSTNSGY-----NKELEFLKNAKESLSKKIEKN 202  
  
QY 112 RLYQIK-----IQYQRENTPEKGLDPK---LYW---TDSQN-----KKEVSSNUL 151  
DB 203 INLLKCDSDNDNLEFKKNVHLNNDYRSEDLFHDKNRVNNQNLKDIKYIYHDDNI 262  
  
QY 152 QLPKQKSGNSRKKRSTAGPTVPDRDNDGIPDSLEVEGYTDV-KNKRFTLSWISNI 210  
DB 263 MI-----EENNDRNDKCNFKNTNPNVERHIIIVDTKNNGNNNNINSNFDN---ISNI 315  
  
QY 211 HEKGLTKYKSPKSTASDP-----YSDFE-KVTGRIDKXVSPPEARHPLVAAYPIVHV- 264  
DB 316 NKK-----ISYPINMYSNSEDKTLNQINLWDSILSSPSLKNAYSFYLS 360  
  
QY 265 ---DMENIILSKNEDQSTQNTD--SETRTISKNTSTSRTHTEVHGNAEVHA--SFFDIG 317  
DB 361 EKNSNNGIKNHSHKNNRYRVIDEDINISFESNISLHDSMDNNEPNSLCSF--- 416  
  
QY 318 GVSAGFSNSSTVAIDHSLSLAGERTWAETVGL-----NTADTARLNANIRY--- 366  
DB 417 -----NSSKCVXDEMITQVGNRKVKYKVTMDANDNMKNSTEQHSINMESHFKNQ 468  
  
QY 367 -VNTGTAPIYVLPVTLVLGKQ-----TLATIKAKE--NOL-SQILAPNN--- 409  
DB 469 KINSKEDANDNLSHILNNKGIGQVNSLSDYLSIKKQESNNNSNNEALNNIV 528  
  
QY 410 -----YYPSSKNLAPI-----ALNAQKDFSTP--- 431  
DB 529 TNNNGSSNNKHSNVYKTSQYVYNNPNDINNHHQNLNLSYMKNSNNLNTSGFKKIPGN 588  
  
QY 432 -----ITWYNOFLEKTKQLRLTDQVYGNIAFYNPENGRVVDGTSNWSVLPQIQ 485  
DB 589 KNIISIDFDNIFKVIKENVYVKNQESNINHGVKNTNDEINIKN-----NNIENNTQ 644  
  
QY 486 ETTAR-----IIFNGKOLNVERRIAANVSDPLETTK----- 518  
DB 645 NITCNPNFINTDIIINKKIKIYKKIDSILSLDLNLSLKNLSNLSIIMDRYTKNYBEK 704  
  
QY 519 ---PDMTLKEAL-----KIAFGNEPNGNLQYQKGDITEFDFNFDQTS 559  
DB 705 FLDDVILDDSIATSNELQHSNYTTNHFIDNNNNNNNNINQKEDLFQNDY---KET 761  
  
QY 560 QNIKNQALAEALNATNIYTVLDKIKLNAMNIIIRDKRPYDRNIAVGADESVVYKEAHREV 619  
DB 762 YNNNIMSENNAENLF-----KISYSCNDLVLGKNEIILD--RNV 800  
  
QY 620 INSGTEGLLINDKIRKILSGVIVEIEDTEGLKE-----VINDRYDMLNIS---L 668  
DB 801 ENSKTE-----QVEEYQNKEDIKLYHKDDNIIIDNNYNDVNIKKDCHL 844  
  
QY 669 RDQKTFIDFKYND---KLPL-----YISNPNYKN-VYAVTKENTIIINPSEN 713  
DB 845 KMDNQDNINKQSDSKKKNPINNECNVIHKNQKFNKINEIDHFKENTILNYSKN 899  
  
RESULT 12  
Q9PQ08 PRELIMINARY; PRT; 4688 AA.  
AC Q9PQ08;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DR 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein UU482.

GN UU482.  
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.  
OX NCBI\_TaxID=134821;  
RW [1]  
RC SEQUENCE FROM N.A.  
PC STRAIN=Serovar 3;  
RX MEDLINE=20500219; PubMed=11048724;  
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
RA Casell G.H.;  
RT "The complete sequence of the mucosal pathogen Ureaplasma  
urealyticum";  
RL Nature 407:757-762(2000).  
DR EMBL; AE002145; AAF30894.1; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 4688 AA; 534880 MW; B53ABFAFEE1997E CRC64;  
  
Query Match 5.4%; Score 205; DB 16; Length 4688;  
Best Local Similarity 21.3%; Pred. No. 0.71;  
Matches 174; Conservative 132; Mismatches 324; Indels 186; Gaps 40;  
  
QY 19 LLGYFSDLPQAPMVVSTTQGLIPSELENIENQYFQSAIWSGFIKVKKDEYV 78  
DB 3699 LVDVYILD-NIHQNIIDETRIKFDKHNV-SKEIENFGVTMISKHGWNKSGPTDTTANFEFK 3756  
  
QY 79 FATSDADNHTVMWDDQEVINKASNSKIRLEKGRLYQIKIQYORENPTEKGLDFKLYWTD 138  
DB 3757 IETQ-----DDNDVLNNIDATVKFDEHNNIKQIVRIKEN-----ND 3795  
  
QY 139 SQNKKEVSSDNLQLPELKOKSN-----SRKKSTAGPTVPRD----- 179  
DB 3796 WLKGGQI---DNLN-PETKYKLENIELSKPLKTHLNSINDKENISLITETGNPVLKV 3851  
  
QY 180 ---NDGIPDSLEVEGYTDVKNKRTFLSPWISNIEKKGLTKYKSPKMWSTASDPYS 235  
DB 3852 IQQNTDINTQTIINTVTLGVSNSK-YNGRQIKVYKDNVNNVYESS---LITLQKQND 3907  
  
QY 236 FEKVTGRIDKNSPEARHPLVAAYPIVHVDMENIILSKN-EDOSTQNTQSET-RTISKNT 293  
DB 3908 YQLLSNLNSN-----REYRFEKIEINISNTNPFEDLEKLGVSNTFITQKNT 3957  
  
QY 294 STSEHTS-EVHGNAEVHASP-----FDIGGSVSAGPS-----SNS 335  
DB 3958 TVQWDSATITGTVGVNPNFKIKSEDKILENNQVAVPAPKETRTDINTWLTQYTRPLK 4017  
  
QY 336 HSLSLAGERTWAETMGLNT---ADTARLNANIRYVNTGTAPIYVNLPTSLVIGKQNTLA 392  
DB 4018 DVTSDPKEGTWAHDLNSNVNFKETTYKLVKIQVKNKPKAKNNNSNNVILDTNISI 4077  
  
QY 393 -----TIKAKENOLSOILAPNNYPSKNLAPALN-AQKDFSTPTWNY--NQFLELE 443  
DB 4078 NSNYEFTTKVGDHKLINITSNNVNTNSQTINFTLSGVKKSWSGKKIKLSYKNDTSESI 4137  
  
QY 444 KTKQLRLTDQVYGNIAFYNPENGR-----VRVDTGSNMSEVLPIQIQTETARIIFNGKD 497  
DB 4138 HTNEVLIESNKTQYNILNLLKRNRTYTLIDVKLIDNNVSDFPKEGNLTNSFITRTSA 4197  
  
QY 498 LNLVERRIAAVNSDPLETT-----KPDWTLKEA-----LKIAFGNEPNGNLQ 541  
DB 4198 INVLNIEISNRASTNLKSTIIKINLNDPNVLRDKQDQATVYGNKQKMGFTVSGNIK 4257  
  
QY 542 YQKGDITEFDFNFDQTSQNIK-NQLAELNATNIYTVLDKIKLNAMNIIIRDKRPYD- 599  
DB 4258 YLTATLDVLPN-DKVNIVNISFNKPSIAAEN-----IGDKSNII-----YNNDS 4304  
  
QY 600 -----RNIIAVGA---DESVVKEAHREVINSSTEGLLINDKIRKILSGVIVEIEDTE 650  
DB 4305 IPKLEINNDIIVGPIKKEIIVKNAOK--NNIDVDLGLQINPKIAHNL-FAIAFKSTN 4361  
  
QY 651 GLKEVIND--RYDMLNSSL-RQDGKTFIDF-----KKY----- 681  
DB 4362 -----NDIETVINGSSLVNDGKTSIRFTLNNLKANKLYSLVDVYVLVNNNSNTIVE 4415



Db 1922 EDIIQERFHNHNTYEMKNKCSYDSCNIIKINYDELNDSTQTKELNEGKSNNGKAEAW 1981  
Qy 106 - IRLKGLRYL-QIKIOYRENPTEKGLDFKLYWTDSONKEVEISSDNLQLPKQKSSNS 163  
Db 1982 IIDIKNNETYPIKIDKEKNEDEK--KNKMYKNKNDKNIMKSSN--KSMKKNKNS 2036  
Qy 164 RKKSTSGAPVPORDNDGIPDSLEVEGYTVVQNKETFLSPWISNHEKGLTKYKSSP 223  
Db 2037 NKMKH-----IFLSVNNKG-----NKSINKKYENNIN-----KONK 2069  
Qy 224 EKWSTASDPYGF--EKVTGRI-DQVSPPEARHPVAAPIVHVHME-----NIIIS 272  
Db 2070 DKLAILVNSIKLVQSKIKQELSNKNSKD-----ILNFEITIKKSKKETKNTNTNN 2124  
Qy 273 KNEDOSTONTSETRTISKNVTSRTHSEV-HGNAEVHASFPFIDGVSNG-PSNSNS 330  
Db 2125 KUNNNNDNNINFERINDINSYVNLNNKVPYDNINI-----DEGNYINYGPYAPDGS 2179  
Qy 331 TVAIHSLSLAGERTWAETMGLNTA-----DTARLNA-----NIRVNTGTAPIYVNL 378  
Db 2180 TI-----YVWVNIDTNYMYNKYFDSKKNINHHVLLNVPVNL-----NDL 2221  
Qy 379 PTTSLVLGKNQFATIKAKENOLSO-----ILAPNYPSPKLAIALNAQKDFSTPTIM 434  
Db 2222 LNNVILNQNMMNLENINTIGSVFVTCDFY--ANNIKSYLD-----PNLQ 2271  
Qy 435 NYNQLELEKTKQLRLDT-----DOVYGNIAFYNFENGVRVDT-----GSN 476  
Db 2272 NNNYFNDMQLNNHNNLNNSINVIDQNYLSYNNLNCINGNINKDTCKDIVIGPNNIN 2331  
Qy 477 WSEVLPQIQTETARIIFNGKDLNVERRIAANVPDPLE-----TTKPDMLKEALK 528  
Db 2332 QNGI-PTI-ELDDTILKN--DVNLLNNNNVNVHSVYVEMLNNIQNVAQKLYNDIQENLH 2387  
Qy 529 IAFGP--NEPNGLOYQKGD-----ITEFDNFQDQTSQNKQNLAEALNATNIVTV 577  
Db 2388 ITNSLHNNNNNNNNMLGGDVLSNVLNFNSFNPNWNAIYNNNNNNNNNNCINTNL 2447  
Qy 578 LDKIKLNAK-----MNLIRDKRPHYDRNNAIAGADESVVKAHREVINSSTEGLLNID 632  
Db 2448 LDRKEINVKEEINMNSVLEDTNKEGND-----MKDNKYIINK--ENNVAQND 2496  
Qy 633 KDRKLSGYVIEDETEGLKEVINDRYDMLISSLRQDKTFIDPKYNDKLPYISNP 692  
Db 2497 NNIDIIISNGQNHNTANTWKEKRNLRDDDDIYKQMKRNSLDFKKLDSE-----KK 2550  
Qy 693 NYKVV--YAVTKENT--IINPSENGDTSTNGIKKILIFSKGYE 733  
Db 2551 NKSNIIEGARKKMTNTQLNYSN--RKGRLDMSTYADKYLE 2593  
PRELIMINARY; PRT; 1404 AA.  
Q8RJN9  
AC Q8RJN9;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Variable membrane protein precursor.  
GN VMP.  
OS Mycoplasma hominis.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2098;  
RN [1]  
RA STRAIN=132;  
RC Boesen T.;  
RT "Gene and Protein Structure of the Mycoplasma hominis Vaa adhesin.";  
RL Thesis (2001), Department of Department of Medical Microbiology and.  
DR EMBL; AJ416752; CAC95143.1; --  
DR InterPro; IPR006864; LMP.  
DR Pfam; PF04776; LMP; 3.

KW Signal. 1 24 POTENTIAL.  
FT SIGNAL 25 1404 VARIABLE MEMBRANE PROTEIN.  
SQ SEQUENCE 1404 AA; 160137 MW; 2C03F66B3473CAB CRC64;  
Query Match 5.1%; Score 194; DB 2; Length 1404;  
Best Local Similarity 19.7%; Pred. NO. 0.5;  
Matches 158; Conservative 124; Mismatches 282; Indels 240; Gaps 34;  
Qy 47 SELENIPEHCYFOSATSWSGFIKYSDEYTFATSNHVTMMVDDQ-----EVIN 98  
Db 312 NSSIEDIQNTKDIKAIES-----LTNKINDQKQKMLNEVIN 351  
Qy 99 KASNSNKIRLEK-GRLYQIKIQORE-----NPTKEGLDFKLYWTDSON 141  
Db 352 KAKELVKLVDSDEIQAKTQDQEIQKASQVASNDDTKAINSSKTSIDAKI--TDITK 409  
Qy 142 KKEVI--SSNLIQIPKQKSSNSRK--KSTSGAPVPORDNDGIPDSLEVEGYTVDVKN 198  
Db 410 KLEAFNATKXLEFTKLOETRSNIDKPLTPEVKANPYTYLVN-----LEVAK 457  
Qy 199 KRTFLSPWISNHEKGLTKYKSS-PEKWTASDPSYDFEKTG-----RI 243  
Db 458 K---AKEWVSSESNKSDIVAANNEKQAFQTAQSSKNDAKTSNEBKAKLSASNAKKL 514  
Qy 244 DKVVS-----PEARHPIVAAPIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSR 297  
Db 515 DKNLTDSDGEIQAKAELAB-----EVEKANOAITSNNTKEIQN--SNTSLNK-----561  
Qy 298 THPSEVHG-----NAEVHASFPFIDGVSVA-----GFSNSNSTVAIDHSLSLAGERTWA 347  
Db 562 --ISEVKNKLDKFSNEKEAEFNKLEASRAIKEFINEENTNPNYALIQKLOAKLDKNS 619  
Qy 348 ETWGLNTADTARLNANIR-----YVNTGTAPIYVNLPTT--SLVAGKN-----388  
Db 620 ITKSSNKSDDIATNQALQELAJEKTESANSQNAVKNLTNETIGKAKELDKNLTDS 679  
Qy 389 ----OTLATIKAKENOLSOILAPNYPSPKLAIALNAQKDFSTPTITMYNYNQFLELEK 444  
Db 680 GETQQAQKAEITNEIEKANQTIASNTNALMENSNTSLNKISEVQNKLDKFNNDKKAENK 739  
Qy 445 TKQLRLDT-----QVYGNIAFYN-FENGVRVDTGSNWS 478  
Db 740 LQELKNKIDDFEKKKNKNEIYKRFNLDELINKSVQIKNSIGSNSSNKKDIIVANKKMQ 799  
Qy 479 EVLPQIQTETARI-----IFN--GKDLNVERRIAANVPSPLETTPKPDMLKEALKI 529  
Db 800 DALNELQAKWAEIHUKTFQEFNEHKELENLIKKEDAKEVGTDEANTAITNNDVKENSSI 859  
Qy 530 AFGFNEPNGLOYQKGDITBDFNFDQTS---QNTKQQLA-ELNATNIVTVLDKIKLNA 585  
Db 860 -----BEITKATKALDEAKSKLDQKINTQKATELSLN-----ESKE 896  
Qy 586 KWNILI-----RDKRPHYDRNNAIAGADESV--VKEAHEVINSSTEGLLN 630  
Db 897 KLNLIITSSSNQVSAEISKAKVLEINLNSLNDSSIKSLKEATQKIDAEQ-LTKE 955  
Qy 631 IDK-----DIRKILSGYIVE-----TEDTEGLKEVINDRYDML 663  
Db 956 IEKAKIEKTDLKRFNEVKKSLEDLIKDDDAIQVGTDDAPKLEDDNNNINE--NSSIEEII 1014  
Qy 664 NISSLRQDKTFIDPKYNDKLP 687  
Db 1015 NATKLEDGKSKLDKIKTKKQPL 1038  
Search completed: May 3, 2004, 19:39:56  
Job time : 39.6261 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:26:27 ; Search time 49.1208 Seconds  
(without alignments)  
4227.791 Million cell updates/sec

Title: US-09-848-909A-10  
Perfect score: 3774  
Sequence: 1 EVKQENLNESSSQGLL.....TSTNGIKKILFSKKGYBIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq.29Jan04:\*

- 1: Genesep1980s.\*
- 2: Genesep1980s.\*
- 3: Genesep2000s.\*
- 4: Genesep2000s.\*
- 5: Genesep2000s.\*
- 6: Genesep2000s.\*
- 7: Genesep2000s.\*
- 8: Genesep2000s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3774	100.0	735	5 AAMS1493	Aams1493 Anthrax P
2	3768	99.8	735	5 AAMS1485	Aams1485 Anthrax P
3	3767	99.8	735	5 AAMS1491	Aams1491 Anthrax P
4	3766	99.8	735	5 AAMS1495	Aams1495 Anthrax P
5	3763	99.7	735	5 AAMS1490	Aams1490 Anthrax P
6	3762	99.7	735	5 AAMS1487	Aams1487 Anthrax P
7	3762	99.7	735	5 AAMS1489	Aams1489 Anthrax P
8	3761	99.7	735	5 AAMS1479	Aams1479 Anthrax P
9	3761	99.7	735	5 AAMS1483	Aams1483 Anthrax P
10	3761	99.7	735	5 AAMS1488	Aams1488 Anthrax P
11	3761	99.7	735	5 AAMS1494	Aams1494 Anthrax P
12	3761	99.7	736	3 AAY56959	Aay56959 B. anthra
13	3761	99.7	763	3 AAY56960	Aay56960 B. anthra
14	3761	99.7	764	3 AAY56958	Aay56958 B. anthra
15	3761	99.7	764	4 AAB47306	Aab47306 Wild type
16	3761	99.7	857	7 ADE65872	Ade65872 Bacillus
17	3760	99.6	735	5 AAMS1484	Aams1484 Anthrax P
18	3759	99.6	735	5 AAMS1486	Aams1486 Anthrax P
19	3758	99.6	735	5 AAMS1489	Aams1489 Anthrax P
20	3758	99.6	764	6 AAE18289	Aae18289 Bacillus
21	3753	99.4	735	6 AAE35717	Aae35717 Bacillus
22	3752	99.4	735	5 AAMS1492	Aams1492 Anthrax P
23	3752	99.4	735	5 AAMS1500	Aams1500 Anthrax P
24	3746	99.3	764	6 AAE18288	Aae18288 Anthrax P
25	3716.5	98.5	903	2 AAR60183	Aar60183 PA(1-725)

26	3620	95.9	711	5 AAMS1498	Aams1498 Anthrax P
27	3615	95.8	711	5 AAMS2113	Aams2113 Anthrax P
28	3614	95.8	711	5 AAMS1497	Aams1497 Anthrax P
29	3607	95.6	711	5 AAMS1496	Aams1496 Anthrax P
30	3584	95.0	719	2 AAR60193	Aar60193 Modified
31	3045	80.7	595	5 AAE18288	Aae18288 Bacillus
32	2906	77.0	569	3 AAY56961	Aay56961 B. anthra
33	2496	66.1	487	5 AAE18285	Aae18285 Bacillus
34	2131	57.8	426	5 AAE18287	Aae18287 Bacillus
35	1632	43.2	318	5 AAE18286	Aae18286 Bacillus
36	1338	35.5	258	5 AAE18284	Aae18284 Bacillus
37	1275	33.8	288	5 AAMS0707	Aams0707 Bacillus
38	847.5	22.5	1052	4 AAE07903	Aae07903 C. botuli
39	846	22.4	1032	4 AAE07901	Aae07901 C. botuli
40	842	22.3	1092	4 AAE07900	Aae07900 C. botuli
41	833.5	22.1	1112	4 AAE07902	Aae07902 C. botuli
42	818	21.7	721	6 AAE35719	Aae35719 Clostridi
43	775.5	20.5	880	2 AAW60224	Aaw60224 Bacillus
44	775	20.5	881	3 AAY59277	Aay59277 MIS toxin
45	774	20.5	884	2 AAR91239	Aar91239 B. cereus

## ALIGNMENTS

### RESULT 1

AAMS1493  
ID AAMS1493 standard; protein; 735 AA.

XX AAMS1493;

DT 01-FEB-2002 (first entry)

DE Anthrax PA mutant K397D/D425K.

XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.

XX Bacillus anthracis.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"

FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"

XX WO200182788-A2.

XX 08-NOV-2001.

XX 04-MAY-2001; 2001WO-US014372.

XX 04-MAY-2000; 2000US-0201800P.

XX (HARD ) HARVARD COLLEGE.

XX Collier RJ, Sellman BR;

XX WPI; 2002-017725/02.

PT Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.

XX Claim 4; Page; 77pp; English.

XX The invention relates to antibacterial agents comprising mutant forms of pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants in the B moiety of the pore-forming binary A-B anthrax toxin, where the B moiety is anthrax protective antigen (PA) and using these mutants or compositions of them for protecting against Bacillus anthracis infections in humans, especially as vaccines. Note: The present sequence is not

CC given in the specification but is derived from the Bacillus anthracis  
CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
XX  
SQ Sequence 735 AA;

Query Match 100.0%; Score 3774; DB 5; Length 735;  
Best Local Similarity 100.0%; Pred. No. 4.3e-241;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLLYGFFSDFLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSQGLLYGFFSDFLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKSDGYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDGYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180

QY 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWKSTASDPYDFEKT 240  
Db 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWKSTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480  
Db 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKALKIAFGFNPNGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERIAAVNPSPDLETTKPDMTLKALKIAFGFNPNGNL 540

QY 541 QYQCKDITEPFDNFQDQTSQNIKNQLAELNATNIYTLVDKIKLNKQNTLIRDKRPHYDR 600  
Db 541 QYQCKDITEPFDNFQDQTSQNIKNQLAELNATNIYTLVDKIKLNKQNTLIRDKRPHYDR 600

QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIKILSGYIVETEDTEGLKEVINDRY 660  
Db 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKDIKILSGYIVETEDTEGLKEVINDRY 660

QY 661 DMLNISLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISLRQDGKTFIDFKYNDKPLIYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720

QY 721 IKKILFSSKGYEIG 735  
Db 721 IKKILFSSKGYEIG 735

RESULT 2  
AAM51485  
ID AAM51485 standard; protein; 735 AA.  
XX  
AC AAM51485;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Anthrax PA mutant K397D.  
XX  
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;

KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX Bacillus anthracis.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 397 /note= "wild-type Lys substituted by Asp"  
XX  
XX WO200182788-A2.  
XX 08-NOV-2001.  
XX 04-MAY-2001; 2001WO-US014372.  
XX 04-MAY-2000; 2000US-0201800P.  
XX (HARD ) HARVARD COLLEGE.  
XX Cellier RJ; Sellman BR;  
XX WPI; 2002-017725/02.  
XX Protecting humans against anthrax using mutant B groups (anthrax protective antigens) of the pore-forming binary A-B toxin of Bacillus anthracis.  
XX Claim 4; Page; 77pp; English.  
XX The invention relates to antibacterial agents comprising mutant forms of pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants in the B moiety of the pore-forming binary A-B anthrax toxin, where the B moiety is anthrax protective antigen (PA) and using these mutants or compositions of them for protecting against Bacillus anthracis infections in humans, especially as vaccines. Note: The present sequence is not given in the specification but is derived from the Bacillus anthracis wild-type PA protein sequence shown in figure 13 (AAM51483)

QY 1 EVKQENRLNESSSQGLLYGFFSDFLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60  
Db 1 EVKQENRLNESSSQGLLYGFFSDFLNFOAPMVVTSSTTGDLSIPSELENIPSENQYF 60

QY 61 QSAIWSGFIKVKSDGYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSDGYTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKQKSNRKRSTAGTVPDRDN 180

QY 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWKSTASDPYDFEKT 240  
Db 181 DGIPOSLVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSPKWKSTASDPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTRTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLSLAGERTWAEWTGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITMNTNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480

Query Match 99.8%; Score 3769; DB 5; Length 735;  
Best Local Similarity 99.9%; Pred. No. 1.1e-240;  
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;



```
Db 421 LNAQDFSSPTITMNYNQFLELEKTKQLRLDQVIGNIATYFENGVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKKNVILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKKNVILIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 3
AAM51491
ID AAM51491 standard; protein; 735 AA.
AC AAM51491;
XX
DT 01-FEB-2002 (first entry)
DE Anthrax PA mutant D425K.
KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.
XX
OS Bacillus anthracis.
XX Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 425
FT /note= "Wild-type Asp substituted by Lys"
XX
PN WO200182788-A2.
XX
XX 08-NOV-2001.
PD
XX PF 04-MAY-2001; 2001WO-US014372.
XX
XX PR 04-MAY-2000; 2000US-0201800P.
XX
XX PA (HARD ) HARVARD COLLEGE.
XX
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX
XX Protecting humans against anthrax using mutant B groups (anthrax
XX protective antigens) of the pore-forming binary A-B toxin of Bacillus
XX anthracis.
XX
XX Claim 4; Page; 77pp; English.
XX
XX The invention relates to antibacterial agents comprising mutant forms of
XX pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants
XX in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
XX moiety is anthrax protective antigen (PA) and using these mutants or
XX compositions of them for protecting against Bacillus anthracis infections
XX in humans, especially as vaccines. Note: The present sequence is not
XX given in the specification but is derived from the Bacillus anthracis
XX wild-type PA protein sequence shown in figure 13 (AAM51483)
XX
```

```
SQ Sequence 735 AA;
Query Match 99.8%; Score 3767; DB 5; Length 735;
Best Local Similarity 99.9%; Pred. No. 1.2e-240;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 EVKQENRLLNESESSQGLLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENYF 60
Db 1 EVKQENRLLNESESSQGLLGYFSDLNFAQPMVVTSTTGDLSIPSSSELENIPSENYF 60
Qy 61 QSAISGFIKVKSDYEYFATISADNHVTWVDDQVINKASNNKIRLEKGLYQIKY 120
Db 61 QSAISGFIKVKSDYEYFATISADNHVTWVDDQVINKASNNKIRLEKGLYQIKY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKKEVISSDNQLQPELKOKSNSRKRKSTAGPTVPDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEVISSDNQLQPELKOKSNSRKRKSTAGPTVPDRDN 180
Qy 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWKSTASDPYDFEKT 240
Qy 241 GRIDKNVSPERHPLVAAPIVHVDMENIILSKNDQSTQNTDSTRIISKNTSRTHT 300
Db 241 GRIDKNVSPERHPLVAAPIVHVDMENIILSKNDQSTQNTDSTRIISKNTSRTHT 300
Qy 301 SEVHGNAEVHAGFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAGFFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSTPTITMNYNQFLELEKTKQLRLDQVIGNIATYFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSTPTITMNYNQFLELEKTKQLRLDQVIGNIATYFENGVRVDTGSNWSEV 480
Qy 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPENG 540
Db 481 LPOIQTETARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPENG 540
Qy 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKKNVILIRDKRPHYDR 600
Db 541 QYQKDIITEFDNFDOOTSQNIKNQLAELNATNIYTVLDKIKLNKKNVILIRDKRPHYDR 600
Qy 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735
```

RESULT 4  
AAM51495  
ID AAM51495 standard; protein; 735 AA.

XX AC AAM51495;  
XX DT 01-FEB-2002 (first entry)  
XX Anthrax PA mutant K397D/D425K/F427A.  
XX DE Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
XX KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.  
XX OS Bacillus anthracis.

OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 397 /note= "Wild-type Lys substituted by Asp"  
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Lys"  
 FT Misc-difference 427 /note= "Wild-type Phe substituted by Ala"  
 FT  
 XX  
 PN WO200182788-A2.  
 XX  
 XX  
 XX 08-NOV-2001.  
 PD  
 XX  
 XX 04-MAY-2001; 2001WO-US014372.  
 PF  
 XX  
 XX 04-MAY-2000; 2000US-0201800P.  
 PR  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 XX  
 XX Collier RJ, Sellman BR;  
 XX  
 XX WPI; 2002-017725/02.  
 DR  
 XX  
 XX Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.  
 XX  
 XX  
 PS Claim 4; Page; 77pp; English.  
 XX  
 XX The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
 XX  
 XX  
 SQ Sequence 735 AA;

Query Match 99.8%; Score 3766; DB 5; Length 735;  
 Best Local Similarity 99.9%; Pred. No. 1.5e-240;  
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNESSSQGLGYFSDLNQFQAPMVVTSSTTGLSTPSSSELENIPSENOYF 60  
 DB 1 EVKQENRLNESSSQGLGYFSDLNQFQAPMVVTSSTTGLSTPSSSELENIPSENOYF 60  
 QY 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
 DB 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
 QY 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLOLPKQKSSNSRKRSTASGTPVDRDN 180  
 DB 121 QRENPTKEGLDFKLYWTDSONKEVIVSSDNLOLPKQKSSNSRKRSTASGTPVDRDN 180  
 QY 181 DGIPDSLEVBGYVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASDPYDFEYKT 240  
 DB 181 DGIPDSLEVBGYVDVKNKRTFLSPWISNIHEKKGTLTKYKSSPEKWSASDPYDFEYKT 240  
 QY 241 GRIDKNVSPARHPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
 DB 241 GRIDKNVSPARHPLVAAPVIVHVDMENILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
 QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360  
 DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSSTVAIDHSLAGERTWAETMGLNTADTARL 360  
 QY 361 NANIRYVNTGTAPVYVLPPTSLVGLGNQTLATIKADENQLSQILAPNNYPSKNLAPTA 420  
 DB 361 NANIRYVNTGTAPVYVLPPTSLVGLGNQTLATIKADENQLSQILAPNNYPSKNLAPTA 420

QY 421 LNAQKDFSTPITMNYNQFLELEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480  
 DB 421 LNAQKDFSTPITMNYNQFLELEKTKQLRLDQVYGNIAFYNFENGVRVDTGNNWSEV 480  
 QY 481 LPQIOETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIARFEPNGNL 540  
 DB 481 LPQIOETTARIIFNGKDLNLVERRIAAVNPSDPLETTKPDMTLKEALKIARFEPNGNL 540  
 QY 541 QYQGGKDI TEFFDNFQDQTSQNIKNQLAELNATNIYTVLDKILNAKMNILIRDKRFHYDR 600  
 DB 541 QYQGGKDI TEFFDNFQDQTSQNIKNQLAELNATNIYTVLDKILNAKMNILIRDKRFHYDR 600  
 QY 601 NNIAVGADES VVKEAHREVINSSTGLLNDKDKIRKILSGYIVIEIDTEGKEVINDRY 660  
 DB 601 NNIAVGADES VVKEAHREVINSSTGLLNDKDKIRKILSGYIVIEIDTEGKEVINDRY 660  
 QY 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPYKVNYYAVTKENTIIINPSNGDTSTNG 720  
 DB 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPYKVNYYAVTKENTIIINPSNGDTSTNG 720  
 QY 721 IKKILIFSKGYEIG 735  
 DB 721 IKKILIFSKGYEIG 735  
 RESULT 5  
 AAM51490  
 ID AAM51490 standard; protein; 735 AA.  
 XX  
 XX AAM51490;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Anthrax PA mutant D425E.  
 XX  
 XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutain.  
 XX  
 XX Bacillus anthracis.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Glu"  
 FT  
 XX  
 PN WO200182788-A2.  
 XX  
 PD 08-NOV-2001.  
 XX  
 PF 04-MAY-2001; 2001WO-US014372.  
 XX  
 PR 04-MAY-2000; 2000US-0201800P.  
 XX  
 XX (HARD ) HARVARD COLLEGE.  
 XX  
 XX Collier RJ, Sellman BR;  
 XX  
 XX WPI; 2002-017725/02.  
 DR  
 XX  
 XX Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.  
 XX  
 XX  
 PS Claim 4; Page; 77pp; English.  
 XX  
 XX The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
 CC

```
XX SQ Sequence 735 AA;
Query Match 99.7%; Score 3763; DB 5; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.3e-240;
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYSFSLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQENRLNSESSESSQGLLYGYSFSLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60

QY 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQLPELKQKSSNRKKGSTAGPTVPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQLPELKQKSSNRKKGSTAGPTVPDRN 180

QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 240
DB 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 480

QY 481 LPOIQTETARIIFNGKDLNLFERRIAAANPSDPLETTKPDMTLKEALKIAGFNEPQNL 540
DB 481 LPOIQTETARIIFNGKDLNLFERRIAAANPSDPLETTKPDMTLKEALKIAGFNEPQNL 540

QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKMWILIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKMWILIRDKRPHYDR 600

QY 601 NNIAVGADESUVKEAHREVINSGTGLLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660
DB 601 NNIAVGADESUVKEAHREVINSGTGLLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660

QY 661 DMLNISLRQDGKTFIDFKKYNKPLPLYISNPNYKNVYAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNISLRQDGKTFIDFKKYNKPLPLYISNPNYKNVYAVTKENTIINPSENGDTSTNG 720

QY 721 IKKILFSKKGVEIG 735
DB 721 IKKILFSKKGVEIG 735

RESULT 6
AAMS1487
ID AAMS1487 standard; protein; 735 AA.
XX AC AAMS1487;
XX DT 01-FEB-2002 (first entry)
XX DE Anthrax PA mutant K397Q.
XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;
XX B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; muten.
XX
```

```
OS Bacillus anthracis.
OS Synthetic.
FH Key Location/Qualifiers
FT Misc-difference 397 /note= "Wild-type Lys substituted by Gln"
FT
XX WO200182788-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001WO-US014372.
XX 04-MAY-2000; 2000US-0201800P.
XX (HARD ) HARVARD COLLEGE.
XX Collier RJ, Sellman BR;
XX WPI; 2002-017725/02.
XX Protecting humans against anthrax using mutant B groups (anthrax
PT protective antigens) of the pore-forming binary A-B toxin of Bacillus
PT anthracis.
XX Claim 4; Page; 77pp; English.
XX The invention relates to antibacterial agents comprising mutant forms of
CC pore-forming toxins (AAMS2113 and AAMS1484-AAMS1500), especially mutants
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B
CC moiety is anthrax protective antigen (PA) and using these mutants or
CC compositions of them for protecting against Bacillus anthracis infections
CC in humans, especially as vaccines. Note: The present sequence is not
CC given in the specification but is derived from the Bacillus anthracis
CC wild-type PA protein sequence shown in figure 13 (AAMS1483)
XX SQ Sequence 735 AA;

Query Match 99.7%; Score 3762; DB 5; Length 735;
Best Local Similarity 99.7%; Pred. No. 2.7e-240;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYSFSLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60
DB 1 EVKQENRLNSESSESSQGLLYGYSFSLNFAQPMVVTSTTGDLSIPSELENIPSENOYF 60

QY 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIVKKSDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQLPELKQKSSNRKKGSTAGPTVPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQLPELKQKSSNRKKGSTAGPTVPDRN 180

QY 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 240
DB 181 DGIPLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSTHT 300

QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILLAPNNYPSKNLAPIA 420

QY 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDTPQVYGNATYNFENGVRVDTGNSWSEV 480
```

481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540  
 481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540  
 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYVLDKIKNKAKMILIRDKGFHYDR 600  
 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYVLDKIKNKAKMILIRDKGFHYDR 600  
 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720  
 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720  
 721 IKKILIFSCKGYEIG 735  
 721 IKKILIFSCKGYEIG 735  
 RESULT 7  
 AAM51489  
 ID AAM51489 standard; protein; 735 AA.  
 XX AC  
 XX AC  
 DT 01-FEB-2002 (first entry)  
 DE Anthrax PA mutant D425N.  
 XX KW Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
 KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutuin.  
 XX OS Bacillus anthracis.  
 OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 425 /note= "Wild-type Asp substituted by Asn"  
 XX WO200182788-A2.  
 XX 08-NOV-2001.  
 XX 04-MAY-2001; 2001WO-US014372.  
 XX 04-MAY-2000; 2000US-0201800P.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Collier RJ, Sellman BR;  
 XX WPI; 2002-017725/02.  
 XX Protecting humans against anthrax using mutant B groups (anthrax  
 PT protective antigens) of the pore-forming binary A-B toxin of Bacillus  
 PT anthracis.  
 XX Claim 4; Page; 77pp; English.  
 CC The invention relates to antibacterial agents comprising mutant forms of  
 CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
 CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
 CC moiety is anthrax protective antigen (PA) and using these mutants or  
 CC compositions of them for protecting against Bacillus anthracis infections  
 CC in humans, especially as vaccines. Note: The present sequence is not  
 CC given in the specification but is derived from the Bacillus anthracis  
 CC wild-type PA protein sequence shown in figure 13 (AAM51483)  
 XX Sequence 735 AA;  
 SQ

Query Match 99.7%; Score 3762; DB 5; Length 735;  
 Best Local similarity 99.7%; Pred. No. 2.7e-240;  
 Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVSTTGGDLSIPSELENIPSENOYF 60  
 DB 1 EVKQENRLNSESSESSQGLLGYYFSDLNFOAPMVVSTTGGDLSIPSELENIPSENOYF 60  
 QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWDQDEVINKASNNKIRLEKGRLYQIKIY 120  
 DB 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTMWDQDEVINKASNNKIRLEKGRLYQIKIY 120  
 QY 121 QSENTEKGLDEKLYWTDSONKKEVSSDNLOLDELKOKSSNRKSTSGAPVPPDRN 180  
 DB 121 QSENTEKGLDEKLYWTDSONKKEVSSDNLOLDELKOKSSNRKSTSGAPVPPDRN 180  
 QY 181 DGIPODSLEVEGYTVDVQKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
 DB 181 DGIPODSLEVEGYTVDVQKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPYDFEKT 240  
 QY 241 GRIDKNVSPERABHPLVAAVPIVHVDMENIILSKNEDOSTQNTDSETRISKTSTSTHT 300  
 DB 241 GRIDKNVSPERABHPLVAAVPIVHVDMENIILSKNEDOSTQNTDSETRISKTSTSTHT 300  
 QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360  
 DB 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLTADTARL 360  
 QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATKADENQLSOILAPNNYPPSKNLAPIA 420  
 DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATKADENQLSOILAPNNYPPSKNLAPIA 420  
 QY 421 LNAQKDFSTPTIMYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNNWSEV 480  
 DB 421 LNAQKDFSTPTIMYNOFLEKTKQLRLDTPQVYGNATYNFENGVRVDTGSNNWSEV 480  
 QY 481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540  
 DB 481 LPOIQTETARIIFNGKDLNLVERRIIAAVNPSPDLETTKPDMTLKEALKIAGFNEPNGNL 540  
 QY 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYVLDKIKNKAKMILIRDKGFHYDR 600  
 DB 541 QYQKDIETEDFNPDQTSQNIKQVLAELNATNIYVLDKIKNKAKMILIRDKGFHYDR 600  
 QY 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
 DB 601 NNIAVGADESUVKEAAREVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660  
 QY 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720  
 DB 661 DMLNLSLRQDGKTFIDFKYNDKLPYISNPYKVVYAVTKENTIIINPSENGDTSTNG 720  
 QY 721 IKKILIFSCKGYEIG 735  
 DB 721 IKKILIFSCKGYEIG 735  
 RESULT 8  
 AAR60179  
 ID AAR60179 standard; protein; 735 AA.  
 XX AC  
 XX AC  
 DT 25-MAR-2003 (revised)  
 DT 03-APR-1995 (first entry)  
 XX DE Protective antigen of Bacillus anthracis.  
 XX KW Anthrax; Bacillus anthracis; fusion protein; protective antigen;  
 KW protective antigen; cell killing; targeting; targeting; pathogen;  
 KW intracellular; HIV; human immunodeficiency virus; toxin.  
 XX OS Bacillus anthracis.

XX WO9418332-A2.  
FN 18-AUG-1994.  
XX 14-FEB-1994; 94WO-US001624.  
XX 12-FEB-1993; 93US-00021601.  
PR 25-JUN-1993; 93US-00082849.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Leppla SH, Klimpel K, Arora N, Singh Y, Nichols PJ;  
XX WPI; 1994-279753/34.  
DR N-PSDB; AAQ70180.  
XX Nucleic acid encoding anthrax toxin fusion protein - useful for  
PT targeting toxin to specific cells, eg for killing tumour cells or HIV-  
PT infected cells.  
XX Disclosure; Page 81-83; 124pp; English.  
XX The sequence encoding the protective antigen of Bacillus anthracis may be  
CC used in the construction of a nucleic acid which encodes a fusion protein  
CC comprising the anthrax protective antigen binding domain of the native  
CC anthrax lethal factor and a sequence encoding an activity inducing domain  
CC of a second protein. The fusion proteins are useful for the specific  
CC killing of tumour cells or the killing of cells infected with  
CC intracellular pathogens, especially HIV. (Updated on 25-MAR-2003 to  
CC correct FN field.)  
XX SQ Sequence 735 AA;  
Query Match 99.7%; Score 3761; DB 2; Length 735;  
Best Local Similarity 99.7%; Pred. No. 3.1e-240;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
QY 61 QSAIWSGFIKVKKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENPTKEGLDFKLYWTDSONKKEVISSDNLQPELKOKSNSRKKRSTAGTVPDRDN 180  
Db 121 QRENPTKEGLDFKLYWTDSONKKEVISSDNLQPELKOKSNSRKKRSTAGTVPDRDN 180  
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKXSTASDPYSDFEVT 240  
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKXSTASDPYSDFEVT 240  
QY 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAPYIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETWGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETWGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVLPPTSILVGLKNOTLATIKADENQLSQILAPNNYFSPKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVLPPTSILVGLKNOTLATIKADENQLSQILAPNNYFSPKNLAPIA 420  
QY 421 LNAQKQFSSPTITWYNNQFLESEKTKQLRLDTPQVYGNATYNFENGRVVRTGNSWSEV 480  
Db 421 LNAQKQFSSPTITWYNNQFLESEKTKQLRLDTPQVYGNATYNFENGRVVRTGNSWSEV 480  
QY 481 LPQIQETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWMTLKEALKIATFAGNEPENG 540  
Db 481 LPQIQETARIIFNGKDLNVERRIAAVNPSPDLETTKPDWMTLKEALKIATFAGNEPENG 540

QY 541 QYQKIDITEFDNFEDQOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDRFRHYDR 600  
Db 541 QYQKIDITEFDNFEDQOTSQNIKNQLAELNATNIYTVLDKIKLNKAKNIIIRDRFRHYDR 600  
QY 601 NNIAGADESVVKEAHREVINSSTEGILLINIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSSTEGILLINIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDPKYNDKPLIYISNPYKAVYAVTKENTINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDPKYNDKPLIYISNPYKAVYAVTKENTINPSENGDTSTNG 720  
QY 721 IKKILIFSCKGYEIG 735  
Db 721 IKKILIFSCKGYEIG 735  
RESULT 9  
AAM51483  
ID AAM51483 standard; protein; 735 AA.  
XX AAM51483;  
XX AC  
XX 01-FEB-2002 (first entry)  
DT Anthrax PA protein.  
DE Anthrax PA protein.  
XX Anthrax; PA; protective antigen; antibacterial; pore-forming toxin;  
KW B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine.  
XX Bacillus anthracis.  
XX OS  
XX WO200182788-A2.  
XX FN  
XX 08-NOV-2001.  
PD  
XX 04-MAY-2001; 2001WO-US014372.  
PF  
XX 04-MAY-2000; 2000US-0201800P.  
PR  
XX (HARD ) HARVARD COLLEGE.  
PA  
XX Collier RJ, Sellman BR;  
PI  
XX WPI; 2002-017725/02.  
XX N-PSDB; AAI99904.  
DR  
XX Protecting humans against anthrax using mutant B groups (anthrax  
PT protective antigens/ of the pore-forming binary A-B toxin of Bacillus  
XX anthracis.  
XX Disclosure; Fig 13; 77pp; English.  
XX The invention relates to antibacterial agents comprising mutant forms of  
CC pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants  
CC in the B moiety of the pore-forming binary A-B anthrax toxin, where the B  
CC moiety is anthrax protective antigen (PA) and using these mutants or  
CC compositions of them for protecting against Bacillus anthracis infections  
CC in humans, especially as vaccines. The present sequence is that of the  
CC anthrax PA protein  
XX SQ Sequence 735 AA;  
Query Match 99.7%; Score 3761; DB 5; Length 735;  
Best Local Similarity 99.7%; Pred. No. 3.1e-240;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
Db 1 EVKQENRLNSESSESSQGLLYGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENYF 60  
QY 61 QSAIWSGFIKVKKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120



QY	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPNKYNVAVTKENTINPSENGDSTNG	720
Db	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPNKYNVAVTKENTINPSENGDSTNG	720
QY	721	IKKILIFSKKGYEIG	735
Db	721	IKKILIFSKKGYEIG	735
RESULT 11			
ID	AAV51494		
XX	AAV51494	standard; protein; 735 AA.	
AC	AAV51494;		
DT	01-FEB-2002	(first entry)	
XX	Anthrax PA mutant K395D/K397D/D425K/D426K.		
XX	Anthrax, PA; protective antigen; antibacterial; pore-forming toxin;		
KW	B moiety; A-B anthrax toxin; Bacillus anthracis; vaccine; mutant; mutein.		
XX	Bacillus anthracis.		
OS	Synthetic.		
XX	Key	Location/Qualifiers	
FT	Misc-difference 395		
FT	/note= "Wild-type Lys substituted by Asp"		
FT	Misc-difference 397		
FT	/note= "Wild-type Lys substituted by Asp"		
FT	Misc-difference 425		
FT	/note= "Wild-type Asp substituted by Lys"		
FT	Misc-difference 426		
FT	/note= "Wild-type Asp substituted by Lys"		
XX	WO200182788-A2.		
PN			
XX			
PD	08-NOV-2001.		
XX			
PF	04-MAY-2001; 2001WO-US014372.		
XX			
PR	04-MAY-2000; 2000US-0201800P.		
XX			
XX	(HARD ) HARVARD COLLEGE.		
PA			
XX	Collier RJ, Sellman BR;		
PI	WPI; 2002-017725/02.		
XX			
DR	Protecting humans against anthrax using mutant B groups (anthrax		
XX	protective antigens) of the pore-forming binary A-B toxin of Bacillus		
PT	anthracis.		
XX			
PS	Claim 4; Page; 77pp; English.		
XX			
CC	The invention relates to antibacterial agents comprising mutant forms of		
CC	pore-forming toxins (AAM52113 and AAM51484-AAM51500), especially mutants		
CC	in the B moiety of the pore-forming binary A-B anthrax toxin, where the B		
CC	moiety is anthrax protective antigen (PA) and using these mutants or		
CC	compositions of them for protecting against Bacillus anthracis infections		
CC	in humans, especially as vaccines. Note: The present sequence is not		
CC	given in the specification but is derived from the Bacillus anthracis		
CC	wild-type PA protein sequence shown in figure 13 (AAM51483).		
XX			
SQ	Sequence 735 AA;		
Query Match			
Best Local Similarity 99.7%; Score 3761; DB 5; Length 735;			
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	EVKQENRLLNESSSSQGLLYGFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60

Db	1	EVKQENRLLNESSSSQGLLYGFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF	60
QY	61	QSAIWSGFTKVYKSDSEYTFATSAADNHVTWVDQVINKASNSNKIRLEKGRLYQIKQY	120
Db	61	QSAIWSGFTKVYKSDSEYTFATSAADNHVTWVDQVINKASNSNKIRLEKGRLYQIKQY	120
QY	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLQQLPELKQKSSNRKGRSTASGTPVDRDN	180
Db	121	QRENPTKGLDFKLYWTDSONKKEVISSDNLQQLPELKQKSSNRKGRSTASGTPVDRDN	180
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKSWSTASDPYDFEYK	240
Db	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSSPEKSWSTASDPYDFEYK	240
QY	241	GRIDKNVSPGARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTTSKNTSRTHT	300
Db	241	GRIDKNVSPGARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTTSKNTSRTHT	300
QY	301	SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
QY	361	NANIRVYNTGTAPIYNNVLPPTSILVGNQTLATIKADENQLSQILAPNNYYPKSLAPIA	420
Db	361	NANIRVYNTGTAPIYNNVLPPTSILVGNQTLATIKADENQLSQILAPNNYYPKSLAPIA	420
QY	421	LNAQKDFSTPTMTNMYNQLELEKTKQLRLDQVYGNATYTNFNGRVRVDTGSNWSEV	480
Db	421	LNAQKDFSTPTMTNMYNQLELEKTKQLRLDQVYGNATYTNFNGRVRVDTGSNWSEV	480
QY	481	LPOIQETTRIIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAGFNEPENGNL	540
Db	481	LPOIQETTRIIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAGFNEPENGNL	540
QY	541	QYQKDIITEFDNFDDQTSQNIKNQAEALNATNYTVLDKIKLNAQONLILRDKRHYDR	600
Db	541	QYQKDIITEFDNFDDQTSQNIKNQAEALNATNYTVLDKIKLNAQONLILRDKRHYDR	600
QY	601	NNIAGADESVKAEHREVNINSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY	660
Db	601	NNIAGADESVKAEHREVNINSTEGLLNIDKIDKILSGYIVETEDTEGLKEVINDRY	660
QY	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPNKYNVAVTKENTINPSENGDSTNG	720
Db	661	DMLNISSLRODQKTFIDFKYNDKLPYISNPNKYNVAVTKENTINPSENGDSTNG	720
QY	721	IKKILIFSKKGYEIG	735
Db	721	IKKILIFSKKGYEIG	735
RESULT 12			
ID	AAV56959	standard; protein; 736 AA.	
XX	AAV56959;		
DT	25-APR-2000	(first entry)	
XX	B. anthracis MAT-PA protein.		
XX	Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;		
KW	tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.		
XX	Bacillus anthracis.		
XX	WO200002522-A2.		
XX	20-JAN-2000.		
XX	09-JUL-1999; 99WO-US015568.		
XX	10-JUL-1998; 98US-0092416P.		

XX (USMB-) US MEDICAL RES INST INFECTIOUS DISEASES.  
XX Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;  
XX WPI; 2000-182165/16.  
XX N-PSDB; AA256875.  
XX Recombinant DNA construct useful as vaccines for anthrax, in producing  
XX host cells for analyzing the drugs and agents inhibiting anthrax.  
XX Disclosure; Page 34; 35pp; English.  
XX The invention provides a recombinant DNA construct that comprises a  
XX vector and at least one nucleic acid (or its fragment) encoding a  
XX combination of Bacillus anthracis proteins, selected from protective  
XX antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA  
XX with its secretory signals replaced with those of tissue plasminogen  
XX activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine  
XX for anthrax and in producing infectious alpha virus particles. These  
XX particles, expressing the B. anthracis proteins are useful also as  
XX vaccines for anthrax. Host cells transformed with the construct are  
XX useful for analyzing the effectiveness of drugs and agents that inhibit  
XX anthrax or B. anthracis proteins. The present sequence represents a B.  
XX anthracis MAT-PA protein  
XX Sequence 736 AA;  
XX  
XX Query Match 99.7%; Score 3761; DB 3; Length 736;  
XX Best Local Similarity 99.7%; Pred. No. 3.1e-240;  
XX Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
XX 2 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 61  
XX 61 QSAIWSGFIKVKKSDRYTATADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
XX 62 QSAIWSGFIKVKKSDRYTATADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 121  
XX 121 QRENTEKGLDFKLYWTDQNKKEVSSDNLQPELKQSSNSRKRKSTASGTPVDRDN 180  
XX 122 QRENTEKGLDFKLYWTDQNKKEVSSDNLQPELKQSSNSRKRKSTASGTPVDRDN 181  
XX 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPFKWSTASDPYSFEKVT 240  
XX 182 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPFKWSTASDPYSFEKVT 241  
XX 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNSTSRHT 300  
XX 242 GRIDKNVSPPEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRISKNSTSRHT 301  
XX 301 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
XX 302 SEVHGNAEVAHSPFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 361  
XX 361 NANIRVYNTGTAPIYNNVPTSLVLGKNOTLATIKADENQLSQILLAPNNYPSKNLPIA 420  
XX 362 NANIRVYNTGTAPIYNNVPTSLVLGKNOTLATIKADENQLSQILLAPNNYPSKNLPIA 421  
XX 421 LNAQKDFSPPTMTNQNQFLEKTKQLRLDQVYGNATYNTFNGRVRVDTGNSWSEV 480  
XX 422 LNAQKDFSPPTMTNQNQFLEKTKQLRLDQVYGNATYNTFNGRVRVDTGNSWSEV 481  
XX 481 LPOIQETTARIIPNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATGFNPNGL 540  
XX 482 LPOIQETTARIIPNGKDLNLVERRIAANVPSDPLETTKPDMTLKEALKIATGFNPNGL 541  
XX 541 QYGGKDIPTDFNFDQQTQNTKQNLAEALNATNIVTLVDKIKLNAKNNILIRDKRFHYDR 600  
XX 542 QYGGKDIPTDFNFDQQTQNTKQNLAEALNATNIVTLVDKIKLNAKNNILIRDKRFHYDR 601  
XX 601 NNTAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660

Db 602 NNTAVGADESVMKEAHEVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 661  
Qy 661 DMLNSSLRQDGKTFIDFKKYNDKLPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720  
Db 662 DMLNSSLRQDGKTFIDFKKYNDKLPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 721  
Qy 721 IKKILFESKGYEIG 735  
Db 722 IKKILFESKGYEIG 736  
RESULT 13  
AAAY56960  
ID AAY56960 standard; protein; 763 AA.  
XX AAY56960;  
XX 25-APR-2000 (first entry)  
DE B. anthracis TPA-PA protein.  
XX Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;  
XX tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.  
XX Bacillus anthracis.  
OS  
XX WO200002522-A2.  
XX 20-JAN-2000.  
XX 09-JUL-1999; 99WO-US015568.  
XX 10-JUL-1998; 98US-0092416P.  
XX (USMB-) US MEDICAL RES INST INFECTIOUS DISEASES.  
XX Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;  
XX WPI; 2000-182165/16.  
XX N-PSDB; AA256875.  
XX Recombinant DNA construct useful as vaccines for anthrax, in producing  
XX host cells for analyzing the drugs and agents inhibiting anthrax.  
XX Disclosure; Page 32; 35pp; English.  
XX The invention provides a recombinant DNA construct that comprises a  
XX vector and at least one nucleic acid (or its fragment) encoding a  
XX combination of Bacillus anthracis proteins, selected from protective  
XX antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA  
XX with its secretory signals replaced with those of tissue plasminogen  
XX activator) and PA63 (63 kDa PA). The construct is useful as a DNA vaccine  
XX for anthrax and in producing infectious alpha virus particles. These  
XX particles, expressing the B. anthracis proteins are useful also as  
XX vaccines for anthrax. Host cells transformed with the construct are  
XX useful for analyzing the effectiveness of drugs and agents that inhibit  
XX anthrax or B. anthracis proteins. The present sequence represents a B.  
XX anthracis TPA-PA protein  
XX Sequence 763 AA;  
XX  
XX Query Match 99.7%; Score 3761; DB 3; Length 763;  
XX Best Local Similarity 99.7%; Pred. No. 3.3e-240;  
XX Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
XX  
XX 1 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 60  
XX 29 EVKQENRLNESSSQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENQYF 88  
XX 61 QSAIWSGFIKVKKSDRYTATADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
XX 89 QSAIWSGFIKVKKSDRYTATADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKIY 148



QY	121	QRENPTKGLDFKLYWTDSONKKEVVISDNQLQPELKQKSSNRKKRSTAGPTVPDRN	180
Db	149	QRENPTKGLDFKLYWTDSONKKEVVISDNQLQPELKQKSSNRKKRSTAGPTVPDRN	208
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYDFEYVT	240
Db	209	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYDFEYVT	268
QY	241	GRIDKNSPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT	300
Db	269	GRIDKNSPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT	328
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLMTADTARL	360
Db	329	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLMTADTARL	388
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	420
Db	389	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA	448
QY	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV	480
Db	449	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV	508
QY	481	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG	540
Db	509	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG	568
QY	541	QYQKXDIETEFDFNPDQOSTQNIKNQALNATNIYTVLDKIKLNAKNLIRDKRPHYDR	600
Db	569	QYQKXDIETEFDFNPDQOSTQNIKNQALNATNIYTVLDKIKLNAKNLIRDKRPHYDR	628
QY	601	NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIYVEIEDTEGLKEVINDRY	660
Db	629	NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIYVEIEDTEGLKEVINDRY	688
QY	661	DLNLISSLRQDGKTIDFKKYNKDLPLYSNPKNVYVAVTKENTIINPSENGDTSTNG	720
Db	689	DLNLISSLRQDGKTIDFKKYNKDLPLYSNPKNVYVAVTKENTIINPSENGDTSTNG	748
QY	721	IKKILIFSKKGYEIG 735	
Db	749	IKKILIFSKKGYEIG 763	
RESULT 14			
ID	AA56958	standard; protein; 764 AA.	
AC	AA56958;		
DT	25-APR-2000 (first entry)		
XX	DE	B. anthracis protective antigen (PA) protein.	
XX	KW	Bacillus anthracis protein; protective antigen; PA; MAT-PA; TPA-PA;	
XX	KW	tissue plasminogen activator; PA63; vaccine; anthrax; antibacterial.	
XX	OS	Bacillus anthracis.	
XX	PN	WC0200002522-A2.	
XX	PD	20-JAN-2000.	
XX	PF	09-JUL-1999; 99WO-US015568.	
XX	PR	10-JUL-1998; 98US-0092416P.	
XX	PA	(USMB-) US MEDICAL RES INST INFECTIOUS DISEASES.	
XX	PI	Lee JS, Pushko P, Parker MD, Smith JF, Welkos SL;	
XX	XX		

DR	WPI; 2000-182165/16.	
DR	N-PSDB; AAZ56874.	
XX	Recombinant DNA construct useful as vaccines for anthrax, in producing	
PT	host cells for analyzing the drugs and agents inhibiting anthrax.	
XX	Disclosure; Page 33; 35pp; English.	
XX	The invention provides a recombinant DNA construct that comprises a	
CC	vector and at least one nucleic acid (or its fragment) encoding a	
CC	combination of Bacillus anthracis proteins, selected from protective	
CC	antigen (PA), MAT-PA (PA with its secretory signals removed), TPA-PA (PA	
CC	with its secretory signals replaced with those of tissue plasminogen	
CC	activator) and PA63 (63 Kda PA). The construct is useful as a DNA vaccine	
CC	for anthrax and in producing infectious alpha virus particles. These	
CC	particles, expressing the B. anthracis proteins are useful also as	
CC	vaccines for anthrax. Host cells transformed with the construct are	
CC	useful for analyzing the effectiveness of drugs and agents that inhibit	
CC	anthrax or B. anthracis proteins. The present sequence represents a B.	
CC	anthracis PA protein	
XX	Sequence 764 AA;	
SQ	Query Match 99.7%; Score 3761; DB 3; Length 764;	
	Best Local Similarity 99.7%; Pred. No. 3.3e-240;	
	Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1	EYKQENRLNSESSESSQGLLYFSDLNAPQAPMVVTSSTGDLSTPSSSELENIPSENQYF 60
Db	30	EYKQENRLNSESSESSQGLLYFSDLNAPQAPMVVTSSTGDLSTPSSSELENIPSENQYF 89
QY	61	QSAIWSGFTKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKQY 120
Db	90	QSAIWSGFTKVKKSDEYTFATSDADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKQY 149
QY	121	QRENPTKGLDFKLYWTDSONKKEVVISDNQLQPELKQKSSNRKKRSTAGPTVPDRN 180
Db	150	QRENPTKGLDFKLYWTDSONKKEVVISDNQLQPELKQKSSNRKKRSTAGPTVPDRN 209
QY	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYDFEYVT 240
Db	210	DGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKSTASDPYDFEYVT 269
QY	241	GRIDKNSPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 300
Db	270	GRIDKNSPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNSTSRHT 329
QY	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLMTADTARL 360
Db	330	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLMTADTARL 389
QY	361	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db	390	NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 449
QY	421	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV 480
Db	450	LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNPENGRVVRVDTGSNWSEV 509
QY	481	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 540
Db	510	LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGNEPENG 569
QY	541	QYQKXDIETEFDFNPDQOSTQNIKNQALNATNIYTVLDKIKLNAKNLIRDKRPHYDR 600
Db	570	QYQKXDIETEFDFNPDQOSTQNIKNQALNATNIYTVLDKIKLNAKNLIRDKRPHYDR 629
QY	601	NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIYVEIEDTEGLKEVINDRY 660
Db	630	NNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIYVEIEDTEGLKEVINDRY 689
QY	661	DLNLISSLRQDGKTIDFKKYNKDLPLYSNPKNVYVAVTKENTIINPSENGDTSTNG 720

Db 690 DMLNLSLRQDGKTFIDFKYNDKPLVYISNPNYKNVYAVTKENTIINPSENGDTSTNG 749

Qy 721 IKKILIFSKKGYEIG 735

Db 750 IKKILIFSKKGYEIG 764

RESULT 15

AA847306

ID AA847306 standard; protein; 764 AA.

XX AA847306;

AC 29-AUG-2001 (first entry)

XX Wild type B. anthracis protective antigen.

DE Lethal factor; LF; immunogen; LF4; protective antigen; PA; DNA vaccine;

XX humoral; cell-mediated; immune memory response.

XX Bacillus anthracis.

XX Key Location/Qualifiers

FT Peptide 1..29

FT /label= Signal peptide

FT /note= "Not given in the specification"

FT Protein 30..764

FT /label= PA

FT Peptide 204..764

FT /label= pCPA

XX WO200145639-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-US034912.

XX 22-DEC-1999; 99US-01714599.

XX (CHIS ) UNIV OHIO STATE RES FOUND.

PA (GALL/) GALLOWAY D R.

PA (MAYE/) MATECZUN A J.

XX Galloway DR, Mateczun AJ;

PI WPI: 2001-408540/43.

DR N-PSDB; AAC86016.

XX

PT Protecting animal against lethal infection with Bacillus anthracis, by

PT administering wildtype or mutated form of Bacillus anthracis lethal

PT factor protein or its fragment or a nucleic acid encoding the mutated

PT protein.

XX Claim 5; Fig 2; 33pp; English.

XX

CC This sequence shows the B. anthracis protective antigen (PA). An

CC immunogenic fragment of PA, pCPA, can be used to produce an immune

CC response which protects an animal against lethal infection with Bacillus

CC anthracis. DNA encoding the B. anthracis PA can be used in conjunction

CC with DNA encoding the lethal factor (LF) in a DNA vaccine. Using a DNA

CC vaccine which encodes the mutated LF protein or fragment alone or in

CC combination with a DNA encoding the PA protein or its fragment, both

CC components (humoral and cell-mediated) of the immune system are

CC stimulated, which results in longer term immune memory response. The

CC combined use of a mutated LF and PA gene or their fragments results in a

CC higher level of immune response, as judged by overall serum antibody

CC titers for LF and PA antigens, than the use of either LF or PA genes in

CC separate immunizations

XX

SQ Sequence 764 AA;

Query Match 99.7%; Score 3761; DB 4; Length 764;

Best Local Similarity 99.7%; Pred. No. 3.3e-240;

Matches	733; Conservative	0; Mismatches	2; Indels	0; Gaps
Qy	1	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSTTGDLSIPSELENIPSENYF	60	
Db	30	EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSTTGDLSIPSELENIPSENYF	89	
Qy	61	QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNNKIRLEKGRLYQIKIY	120	
Db	90	QSAIWSGFIKVKSDSEYTFATSDNHVTWVDDQEVINKASNNKIRLEKGRLYQIKIY	149	
Qy	121	QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNKRKSTAGTPVPDRDN	180	
Db	150	QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNKRKSTAGTPVPDRDN	209	
Qy	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYDFEKT	240	
Db	210	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHKKGLTKYKSPKSWSTASDPYDFEKT	269	
Qy	241	GRIDKNVSPKARHPLVAAVPIVHVDNENIILSKNEDQSTQNTDSETRISKNTSRTHT	300	
Db	270	GRIDKNVSPKARHPLVAAVPIVHVDNENIILSKNEDQSTQNTDSETRISKNTSRTHT	329	
Qy	301	SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL	360	
Db	330	SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL	389	
Qy	361	NANIRYVNTGTAPIYNNVPTTSLVLGKQNTLTIKADENQLSQILAPNNYPSKNLAPIA	420	
Db	390	NANIRYVNTGTAPIYNNVPTTSLVLGKQNTLTIKADENQLSQILAPNNYPSKNLAPIA	449	
Qy	421	LNAQKDFSTPITMNYNQFLELEKTKQLRLDQVYGNIAATYNFENGRVVDVTGNSWSEV	480	
Db	450	LNAQKDFSTPITMNYNQFLELEKTKQLRLDQVYGNIAATYNFENGRVVDVTGNSWSEV	509	
Qy	481	LPOIQTETARIIFNGKOLNVERRIAANVPSPLETTKEDMTLKEALKIAGFNPNGNL	540	
Db	510	LPOIQTETARIIFNGKOLNVERRIAANVPSPLETTKEDMTLKEALKIAGFNPNGNL	569	
Qy	541	QYQKDIETFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKPFHYDR	600	
Db	570	QYQKDIETFDNFDOQTQSNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKPFHYDR	629	
Qy	601	NNIAGADESVVKEAHREVINSTEGLLNIDKIRKILSGYVIEIETEGEKEVINDRY	660	
Db	630	NNIAGADESVVKEAHREVINSTEGLLNIDKIRKILSGYVIEIETEGEKEVINDRY	689	
Qy	661	DMLNLSLRQDGKTFIDFKYNDKPLVYISNPNYKNVYAVTKENTIINPSENGDTSTNG	720	
Db	690	DMLNLSLRQDGKTFIDFKYNDKPLVYISNPNYKNVYAVTKENTIINPSENGDTSTNG	749	
Qy	721	IKKILIFSKKGYEIG 735		
Db	750	IKKILIFSKKGYEIG 764		

Search completed: May 3, 2004, 19:35:58

Job time : 52.1208 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:31:43 ; Search time 15.0986 Seconds  
(without alignments)  
2513.152 Million cell updates/sec

Title: US-09-848-909A-10  
Perfect score: 3774  
Sequence: 1 EVKQENRLNSESQGLL.....TSTNGIKILFFSKGYEIG 735  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues  
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pep:\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pep:\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pep:\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pep:\*  
5: /cgn2\_6/prodata/2/iaa/PCFUS COMB.pep:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3761	99.7	735	1 US-08-021-601-4	Sequence 4, Appli
2	3761	99.7	735	1 US-08-082-849B-4	Sequence 4, Appli
3	3761	99.7	735	5 PCT-US94-01624-4	Sequence 4, Appli
4	3716.5	98.5	903	1 US-08-021-601-12	Sequence 12, Appli
5	3716.5	98.5	903	1 US-08-082-849B-12	Sequence 12, Appli
6	3716.5	98.5	903	5 PCT-US94-01624-12	Sequence 12, Appli
7	3584	95.0	719	1 US-08-082-849B-31	Sequence 31, Appli
8	3584	95.0	719	5 PCT-US94-01624-31	Sequence 31, Appli
9	1275	33.8	288	4 US-09-273-839A-8	Sequence 8, Appli
10	775	20.5	881	3 US-08-960-780-32	Sequence 32, Appli
11	775	20.5	881	3 US-09-073-898-32	Sequence 32, Appli
12	775	20.5	881	4 US-09-307-106-8	Sequence 8, Appli
13	775	20.5	881	4 US-09-850-351A-32	Sequence 32, Appli
14	773	20.5	884	1 US-08-471-033-5	Sequence 5, Appli
15	773	20.5	884	2 US-08-471-044-5	Sequence 5, Appli
16	773	20.5	884	2 US-08-463-483A-5	Sequence 5, Appli
17	773	20.5	884	2 US-08-471-046A-5	Sequence 5, Appli
18	773	20.5	884	2 US-08-470-566B-5	Sequence 5, Appli
19	773	20.5	884	2 US-08-469-334-5	Sequence 5, Appli
20	773	20.5	884	3 US-09-300-529-5	Sequence 5, Appli
21	773	20.5	1346	1 US-08-471-033-23	Sequence 23, Appli
22	773	20.5	1346	2 US-08-471-044-23	Sequence 23, Appli
23	773	20.5	1346	2 US-08-463-483A-23	Sequence 23, Appli
24	773	20.5	1346	2 US-08-471-046A-23	Sequence 23, Appli
25	773	20.5	1346	2 US-08-470-566B-23	Sequence 23, Appli
26	773	20.5	1346	2 US-08-469-334-23	Sequence 23, Appli
27	773	20.5	1346	3 US-09-300-529-23	Sequence 23, Appli

28	772	20.5	852	1 US-08-471-033-36	Sequence 36, Appli
29	772	20.5	852	2 US-08-471-044-36	Sequence 36, Appli
30	772	20.5	852	2 US-08-463-483A-36	Sequence 36, Appli
31	772	20.5	852	2 US-08-471-046A-36	Sequence 36, Appli
32	772	20.5	852	2 US-08-470-566B-36	Sequence 36, Appli
33	772	20.5	852	2 US-08-469-334-36	Sequence 36, Appli
34	772	20.5	852	3 US-09-300-529-36	Sequence 36, Appli
35	772	20.5	1338	1 US-08-471-033-50	Sequence 50, Appli
36	772	20.5	1338	2 US-08-471-044-50	Sequence 50, Appli
37	772	20.5	1338	2 US-08-463-483A-50	Sequence 50, Appli
38	772	20.5	1338	2 US-08-471-046A-50	Sequence 50, Appli
39	772	20.5	1338	2 US-08-470-566B-50	Sequence 50, Appli
40	772	20.5	1338	2 US-08-469-334-50	Sequence 50, Appli
41	772	20.5	1338	3 US-09-300-529-50	Sequence 50, Appli
42	762.5	20.2	784	3 US-09-371-913A-7	Sequence 7, Appli
43	762.5	20.2	784	4 US-09-967-805-7	Sequence 48, Appli
44	753.5	20.0	860	4 US-09-307-106-48	Sequence 48, Appli
45	743.5	19.7	834	1 US-08-471-033-21	Sequence 21, Appli

ALIGNMENTS

RESULT 1  
US-08-021-601-4  
; Sequence 4, Application US/08021601  
; Patent No. 5591631  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Nichols, Peter J.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Needle & Rosenberg, P.C.  
; STREET: 133 Carnegie Way, Suite 400  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021.601  
; FILING DATE: 19930212  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 1414.057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-021-601-4

Query Match 99.7%; Score 3761; DB 1; Length 735;  
Best Local Similarity 99.7%; Pred No. 1.2e-254;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESQGLLGYFSDLNFOQPMVVTSTTGDLSIPSELNISENQYF 60  
|||||

Db 1 EVKQENRLNBSSESSQGLGYFSDLNFOAPMVVTSITGDLSPSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKKSDYFATSAADHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDYFATSAADHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENTEKGLDFKLYWTSQNKKEVSSDNIQLPELKQKSSNSKRSKSTAGPTVPDRN 180  
Db 121 QRENTEKGLDFKLYWTSQNKKEVSSDNIQLPELKQKSSNSKRSKSTAGPTVPDRN 180  
Qy 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240  
Db 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240  
Qy 241 GRIDKNVSPERHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTSKNTSTSTHT 300  
Db 241 GRIDKNVSPERHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTSKNTSTSTHT 300  
Qy 301 SEVHGNAEVAHAFDYGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHAFDYGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSTPTMTWYNOFLEKTKQLELDTDOVYGNIAITVNFENGVRVDTGNSNVEV 480  
Db 421 LNAQKDFSTPTMTWYNOFLEKTKQLELDTDOVYGNIAITVNFENGVRVDTGNSNVEV 480  
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540  
Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLRDKRFFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLRDKRFFHYDR 600  
Qy 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660

RESULT 2  
US-082-849B-4  
Sequence 4, Application US/08082849B  
Patent No. 567274  
GENERAL INFORMATION:  
APPLICANT: Leppia, Stephen H.  
APPLICANT: Klimpel, Kurt R.  
APPLICANT: Arora, Naveen  
APPLICANT: Singh, Yogendra  
APPLICANT: Nichols, Peter J.  
TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
TITLE OF INVENTION: Related Methods  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,849B  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/021,601  
FILING DATE: 12-FEB-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Kenneth A.  
REGISTRATION NUMBER: 31,677  
REFERENCE/DOCKET NUMBER: 15280-161-1  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-082-849B-4

Query Match 99.7%; Score 3761; DB 1; Length 735;  
Best Local Similarity 99.7%; Pred. No. 1.2e-254;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 EVKQENRLNBSSESSQGLGYFSDLNFOAPMVVTSITGDLSPSELENIPSENQYF 60  
Db 1 EVKQENRLNBSSESSQGLGYFSDLNFOAPMVVTSITGDLSPSELENIPSENQYF 60  
Qy 61 QSAIWSGFIKVKKSDYFATSAADHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDYFATSAADHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Qy 121 QRENTEKGLDFKLYWTSQNKKEVSSDNIQLPELKQKSSNSKRSKSTAGPTVPDRN 180  
Db 121 QRENTEKGLDFKLYWTSQNKKEVSSDNIQLPELKQKSSNSKRSKSTAGPTVPDRN 180  
Qy 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240  
Db 181 DGIPDSLEVGTVVVKKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240  
Qy 241 GRIDKNVSPERHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTSKNTSTSTHT 300  
Db 241 GRIDKNVSPERHPLVAAPYIVHVDMENILSKNEDQSTQNTDSETRTSKNTSTSTHT 300  
Qy 301 SEVHGNAEVAHAFDYGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHAFDYGVSAGFSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSTPTMTWYNOFLEKTKQLELDTDOVYGNIAITVNFENGVRVDTGNSNVEV 480  
Db 421 LNAQKDFSTPTMTWYNOFLEKTKQLELDTDOVYGNIAITVNFENGVRVDTGNSNVEV 480  
Qy 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540  
Db 481 LPQIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNCNL 540  
Qy 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLRDKRFFHYDR 600  
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAKNILLRDKRFFHYDR 600  
Qy 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTGLKEVINDRY 660

QY 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNKYVAVTKENTINPSENGDTSTNG 720  
DB 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNKYVAVTKENTINPSENGDTSTNG 720  
QY 721 IKKILFSSKGYEIG 735  
DB 721 IKKILFSSKGYEIG 735  
RESULT 3  
PCT-US94-01624-4  
; Sequence 4, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: TOWNSEND and TOWNSEND KHOURIE and CREW  
; STREET: Steuart Street Tower, 20th Floor, One Market  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01624  
; FILING DATE: June 25, 1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-01624-4  
Query Match 99.7%; Score 3761; DB 5; Length 735;  
Best Local Similarity 99.7%; Pred. No. 1.2e-254;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVQENRLNNESSSQGLLYYFSDLPNFPQPMVVTSTTGDLSPSSLENIPSENOYF 60  
DB 1 EVQENRLNNESSSQGLLYYFSDLPNFPQPMVVTSTTGDLSPSSLENIPSENOYF 60  
QY 61 QSAIWGFIKVKSDYVTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKYQ 120  
DB 61 QSAIWGFIKVKSDYVTFATSDADNHVTWVDDQEVINKASNKIRLEKGRLYQIKYQ 120  
QY 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQLPKQSSNRKRSSTAGTVPDRDN 180  
DB 121 QRENPTKGLDFKLYWTDSONKEVSSDNLQLPKQSSNRKRSSTAGTVPDRDN 180  
QY 181 DGIPTDLSLEVEGYVDVKNKRTFSPWISNHEKKGITKYSPEKWTASDPSDFEYVT 240  
DB 181 DGIPTDLSLEVEGYVDVKNKRTFSPWISNHEKKGITKYSPEKWTASDPSDFEYVT 240

QY 241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300  
DB 241 GRIDKNVSEARHPLVAAYPIVHVDMENILSKNEDQSTQNTDSETRTTSKNTSTSRHT 300  
QY 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
DB 301 SEVHGNAEVHASFDDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKAKENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDPSSPTITMNNYQFLEKTKQLRLDDOYVGNATYFNENGRVVRVDTGSNWSEV 480  
DB 421 LNAQKDPSSPTITMNNYQFLEKTKQLRLDDOYVGNATYFNENGRVVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENG 540  
DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLKEALKIAFGNEPENG 540  
QY 541 QYQKDIITBDFNFDQDTSQNTKQNLAEALNATYVTLQKIKLNKQNLILRDKRFHYDR 600  
DB 541 QYQKDIITBDFNFDQDTSQNTKQNLAEALNATYVTLQKIKLNKQNLILRDKRFHYDR 600  
QY 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
DB 601 NNIAVGADSVVKEAHEVINSSTEGLLNIDKIRKILSGYIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNKYVAVTKENTINPSENGDTSTNG 720  
DB 661 DMLNSSLRQDQKTFIDFKYNDKLPYISNPNKYVAVTKENTINPSENGDTSTNG 720  
QY 721 IKKILFSSKGYEIG 735  
DB 721 IKKILFSSKGYEIG 735  
RESULT 4  
US-08-021-601-12  
; Sequence 12, Application US/08021601  
; Patent No. 5591631  
; GENERAL INFORMATION:  
; APPLICANT: Leppla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Nichols, Peter J.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; RELATED METHODS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Needle & Rosenberg, P.C.  
; STREET: 133 Carnegie Way, Suite 400  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/021,601  
; FILING DATE: 19930212  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Spratt, Gwendolyn D.  
; REGISTRATION NUMBER: 36,016  
; REFERENCE/DOCKET NUMBER: 1414.057  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404/688-0770  
; TELEFAX: 404/688-9880

```

; INFORMATION FOR SEQ ID NO: 12:
;
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 903 amino acids
;     TYPE: AMINO ACID
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-021-601-12

```

Query Match 98.5%; Score 3716.5; DB 1; Length 903;  
Best Local Similarity 98.9%; Pred. No. 2e-251;  
Matches 726; Conservative 1; Mismatches 4; Indels 3;

Qy	1	EYKQENRLN	EBSESQGLGYYS	DLNFQAPMV	TSSTTCDLS	ISPSSELENI	PSENQYF	60							
Db	1	EYKQENRLN	EBSESQGLGYYS	DLNFQAPMV	TSSTTCDLS	ISPSSELENI	PSENQYF	60							
Qy	61	QSAIWSGFI	KVKKSD	DEYTFATS	ADNHVTMW	VDDQEVINK	ASNKNIRLEK	GLYQIKIY 120							
Db	61	QSAIWSGFI	KVKKSD	DEYTFATS	ADNHVTMW	VDDQEVINK	ASNKNIRLEK	GLYQIKIY 120							
Qy	121	QRENPTKGL	DFKLYW	TOSQNK	EVISSDNI	LOLP	ELKQKSSNR	KKRSAGTVPYDPRDN 180							
Db	121	QRENPTKGL	DFKLYW	TOSQNK	EVISSDNI	LOLP	ELKQKSSNR	KKRSAGTVPYDPRDN 180							
Qy	181	DGIPDSLE	VEGYTVDV	QKNKRTFL	SPWISNI	HEKKGLTKYK	SSPEKWS	TASDPYSDFEKT 240							
Db	181	DGIPDSLE	VEGYTVDV	QKNKRTFL	SPWISNI	HEKKGLTKYK	SSPEKWS	TASDPYSDFEKT 240							
Qy	241	GRIDKXV	SEAP	HLVAAYPI	VHVDMENI	ILSKNED	OSTONT	USETRTIS	SKTSSRTHT 300						
Db	241	GRIDKXV	SEAP	HLVAAYPI	VHVDMENI	ILSKNED	OSTONT	USETRTIS	SKTSSRTHT 300						
Qy	301	SEVHGNAE	VHASF	FFDIGSV	SAGFSNS	STVAIDHSL	SLAGERT	WAE	TWGLNTADTARL 360						
Db	301	SEVHGNAE	VHASF	FFDIGSV	SAGFSNS	STVAIDHSL	SLAGERT	WAE	TWGLNTADTARL 360						
Qy	361	NANIRVYNT	GTAPIV	NVLPTTSL	VLGKQ	QTUAT	KADENQL	SOILLAP	NNYPSKNI	LAPIA 420					
Db	361	NANIRVYNT	GTAPIV	NVLPTTSL	VLGKQ	QTUAT	KADENQL	SOILLAP	NNYPSKNI	LAPIA 420					
Qy	421	LNAAQDFS	STPTIM	YNQOFLE	LEKTKOL	RLOTDOV	YGNIA	TATYFNGR	VRVD	TGNSWSEV 480					
Db	421	LNAAQDFS	STPTIM	YNQOFLE	LEKTKOL	RLOTDOV	YGNIA	TATYFNGR	VRVD	TGNSWSEV 480					
Qy	481	LPQIOET	TARIIF	NGKOLNL	VERRIAA	VNP	SDPLETT	KPDM	TIKEAL	KIAGFNE	PNGL 540				
Db	481	LPQIOET	TARIIF	NGKOLNL	VERRIAA	VNP	SDPLETT	KPDM	TIKEAL	KIAGFNE	PNGL 540				
Qy	541	QYQKDI	TDFDN	FOOTSQ	NIKNQ	LAELNAT	INTYVLDKI	KNAKM	NIILIR	KRPHYDR 600					
Db	541	QYQKDI	TDFDN	FOOTSQ	NIKNQ	LAELNAT	INTYVLDKI	KNAKM	NIILIR	KRPHYDR 600					
Qy	601	NNI	AVGADES	VVKEA	REVIN	STEG	LLNID	KOIRKIL	SGYI	VEIBDT	EGLKEVIN	DY 660			
Db	601	NNI	AVGADES	VVKEA	REVIN	STEG	LLNID	KOIRKIL	SGYI	VEIBDT	EGLKEVIN	DY 660			
Qy	661	DM	NI	SSLRO	DKGT	FIDFKY	NDKLP	LYTIS	HPN	YKNV	YATV	KENT	TIN	PSENGD	TSTNG 720
Db	661	DM	NI	SSLRO	DKGT	FIDFKY	NDKLP	LYTIS	HPN	YKNV	YATV	KENT	TIN	PSENGD	TSTNG 720
Qy	721	IKKIL	---	IF	SKKG	731									
Db	721	IKKIL	---	IF	SKKG	734									

RESULT 5  
US-082-849B-12  
; Sequence 12, Application US/08092849B  
; Patent No. 5677274  
; GENERAL INFORMATION:  
; APPLICANT: Lebpla, Stephen H.  
; APPLICANT: Klimpel, Kurt R.  
; APPLICANT: Arota, Naveen

APPLICANT: Singh, Yogendra  
 APPLICANT: Nichols, Peter J.  
 TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
 TITLE OF INVENTION: Related Methods  
 NUMBER OF SEQUENCES: 35  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/082,849B  
 FILING DATE: 25-JUN-1993  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/021,601  
 FILING DATE: 12-FEB-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Weber, Kenneth A.  
 REGISTRATION NUMBER: 31,677  
 REFERENCE/DOCKET NUMBER: 15280-161-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 903 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-082-849B-12

Query Match	98.5%; Score 3716.5; DB 1; Length 903;
Best Local Similarity	98.9%; Pred. No. 2e-251;
Matches	726; Conservative 1; Mismatches 4; Indels 3; Gaps 1
Qy	1 EVKQENRLNLESSESSQGLLGYFSDLNQCAPMVVTSITTTGDLSPSSLENIPSENQYF 60
Db	1 EVKQENRLNLESSESSQGLLGYFSDLNQCAPMVVTSITTTGDLSPSSLENIPSENQYF 60
Qy	61 QSAIWSGFIKVKKSDEYTFATSDADNHYTMVMDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db	61 QSAIWSGFIKVKKSDEYTFATSDADNHYTMVMDQEVINKASNSNKIRLEKGRLYQIKIY 120
Qy	121 QRENPTKEGLDFKLWYTDSONKKEVISSDNLQLPFLKQKSSNSRKKRSTAGPTVPORDN 180
Db	121 QRENPTKEGLDFKLWYTDSONKKEVISSDNLQLPFLKQKSSNSRKKRSTAGPTVPORDN 180
Qy	181 DGIPTDSLSEVGTYVDVKNKETFSPWISNTHKKGLTKYKSSPEKWTASDPPYSDPEKVT 240
Db	181 DGIPTDSLSEVGTYVDVKNKETFSPWISNTHKKGLTKYKSSPEKWTASDPPYSDPEKVT 240
Qy	241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDOSTONTUSETTTISKNTSTSRTHT 300
Db	241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDOSTONTUSETTTISKNTSTSRTHT 300
Qy	301 SEVHGNAEVHASPPDIGGSVAGSPNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db	301 SEVHGNAEVHASPPDIGGSVAGSPNSNSSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy	361 NANTRYNTGTAPYIYNVLPTSLVLGKNQTLATTIKADENQLSQILAPNNYYPSPKNLAPIA 420
Db	361 NANTRYNTGTAPYIYNVLPTSLVLGKNQTLATTIKADENQLSQILAPNNYYPSPKNLAPIA 420
Qy	421 LNAQKQFSSTPIITNNYNQFLELEKTKQLRLDPTDQVGYGNIATYNFENGVRVYDTGSNNSEV 480

Db 421 LNAQDDFSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNNSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540  
QY 541 QYQGDITEFDNFDOQTSONIKNOQLAELNATNIYVLDKIKINAKNMLIRDKGFHYDR 600  
Db 541 QYQGDITEFDNFDOQTSONIKNOQLAELNATNIYVLDKIKINAKNMLIRDKGFHYDR 600  
QY 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660  
QY 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNKYAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNKYAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKIL---IFSKKG 731  
Db 721 IKKILKVVGLKKG 734

RESULT 6  
PCT-US94-01624-12  
; Sequence 12, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Kimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE and CREW  
; STREET: Steuart Street Tower, 20th Floor, One Market  
; Plaza  
; City: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01624  
; Filing Date: June 25, 1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 903 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-01624-12

Query Match 98.5%; Score 3716.5; DB 5; Length 903;  
Best Local Similarity 98.9%; Pred. No. 2e-251;  
Matches 726; Conservative 1; Mismatches 4; Indels 3; Gaps 1;

QY 1 EVKQENRLNSESSESSQGLLYYFSDLPQAPMVVTSSTTGLSLPSPSELENIIPSENQYF 60  
61 QSAINSGFIKVKKSDSEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
61 QSAINSGFIKVKKSDSEYTFATSAADNHVTMWVDDQEVINKASNSKIRLEKGLYQIKQY 120  
121 QRENPEKGLDFKLYWTDSONKKEVISSDNLOIPELKQKSSNSRKKRSTASAGTVPDRN 180  
121 QRENPEKGLDFKLYWTDSONKKEVISSDNLOIPELKQKSSNSRKKRSTASAGTVPDRN 180  
181 DGIPDLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSPFVKY 240  
181 DGIPDLSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSPFVKY 240  
241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
241 GRIDKNVSEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
301 SEVHGNAEVAHFFDIDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
301 SEVHGNAEVAHFFDIDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
421 LNAQDDFSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNNSEV 480  
421 LNAQDDFSTPTIMYNOFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNNSEV 480  
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540  
481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAFGFNEPNGNL 540  
541 QYQGDITEFDNFDOQTSONIKNOQLAELNATNIYVLDKIKINAKNMLIRDKGFHYDR 600  
541 QYQGDITEFDNFDOQTSONIKNOQLAELNATNIYVLDKIKINAKNMLIRDKGFHYDR 600  
601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660  
601 NNIAGADESVVKEAHREVINSSTEGLLNIDKDIRKILSGYIIEDETEGLKEVINDRY 660  
661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNKYAVTKENTIINPSENGDTSTNG 720  
661 DMLNISSLRQDGKTFIDFKKYNKPLIYISNPNYKVNKYAVTKENTIINPSENGDTSTNG 720  
721 IKKIL---IFSKKG 731  
721 IKKILKVVGLKKG 734

RESULT 7  
US-08-082-849B-31  
; Sequence 31, Application US/08082849B  
; Patent No. 5677274  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Kimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and  
; TITLE OF INVENTION: Related Methods  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; City: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy

```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patettin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,849B
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,601
; FILING DATE: 12-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
; REFERENCE/DOCKET NUMBER: 15280-161-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 719 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-082-849B-31
;
Query Match 95.0%; Score 3584; DB 1; Length 719;
Best Local Similarity 95.3%; Pred. No. 2.6e-242;
Matches 704; Conservative 4; Mismatches 7; Indels 24; Gaps 2;

QY 1 EVKQENRLNSESSESQGLGYVSDLNFOAPWVTTSITGDLSPSSLENIIPSENQYF 60
DB 1 EVKQENRLNSESSESQGLGYVSDLNFOAPWVTTSITGDLSPSSLENIIPSENQYF 60
QY 61 QSAIWSGFIKVKKSDSEYTFATSDAHVITMWYDDQEVINKASNKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKKSDSEYTFATSDAHVITMWYDDQEVINKASNKIRLEKGLYQIKIY 120
QY 121 QRENPTKEGLDKPLWYTSQNKKEVSSDNLQLPBKQKSSNS-----RKRSTSGAPTPV 176
DB 121 QRENPTKEGLDKPLWYTSQNKKEVSSDNLQLPBKQKSSNTATIMMQRGNPLOGTPV 180
QY 177 DRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASTSPYSDP 236
DB 181 DRDNDGIPDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKWSASTSPYSDP 240
QY 237 EKYTGRIKQNSPEARHPLVAAYPLIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTS 296
DB 241 EKYTGRIKQNSPEARHPLVAAYPLIVHVDMENIILSKNEDQSTQNTDSETRISKNTSTS 300
QY 297 RHTTSEVHGNAEVHASFDDIGGSVAGSPNSNSTVAIDHSLAGERTWAETWGLNTAD 356
DB 301 RHTTSEVHGNAEVHASFDDIGGSVAGSPNSNSTVAIDHSLAGERTWAETWGLNTAD 360
QY 357 TARLNANIYVNTGTAPIYNYVLPPTSLVLGKXQTLATIKADENQLQSILAPNNYPSKUL 416
DB 361 TARLNANIYVNTGTAPIYNYVLPPTSLVLGKXQTLATIKAKENQLQSILAPNNYPSKUL 420
QY 417 APALNAQDQFSSTPIITMNYNQFLEKTKQLRLDTPQVGNATYNFENGVRVDVTGSN 476
DB 421 APALNAQDQFSSTPIITWN-----YGNATYNFENGVRVDVTGSN 460
QY 477 WSEVLPOIQETTARIIFNGKQNLNVERRIAANPSDPLETTKPDMTLKALKIAFGNEFP 536
DB 461 WSEVLPOIQETTARIIFNGKQNLNVERRIAANPSDPLETTKPDMTLKALKIAFGNEFP 520
QY 537 NGNLQYQGGKDIETFPNFDQOTSQNIKNQALNATNIYTVLDKIKLNAGWILLRDKRF 596
DB 521 NGNLQYQGGKDIETFPNFDQOTSQNIKNQALNATNIYTVLDKIKLNAGWILLRDKRF 580
QY 597 HYDRNNIAYGADESYYVKAHREVINSSTEGLLINIDKIRKILSGYVIEDETEGLKEYI 656
DB 581 HYDRNNIAYGADESYYVKAHREVINSSTEGLLINIDKIRKILSGYVIEDETEGLKEYI 640

```

657 NDRYDMLNISLSLRQDGKTFIDFKYNDKCLPLYISNPNYKYNVAVTKENTIIINPSENGDT 711  
Db 641 NDRYDMLNISLSLRQDGKTFIDFKYNDKCLPLYISNPNYKYNVAVTKENTIIINPSENGDT 700  
QY 717 STNGIKKILIFSKGYEIG 735  
Db 701 STNGIKKILIFSKGYEIG 719  
RESULT 8  
PCT-US94-01624-31  
; Sequence 31, Application PC/TUS9401624  
; GENERAL INFORMATION:  
; APPLICANT: Leppia, Stephen H.  
; APPLICANT: Kimpel, Kurt R.  
; APPLICANT: Arora, Naveen  
; APPLICANT: Singh, Yogendra  
; APPLICANT: Nichols, Peter J.  
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND  
; TITLE OF INVENTION: RELATED METHODS  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND AND TOWNSEND KHOURIE and CREW  
; STREET: Stewart Street Tower, 20th Floor, One Market  
; STREET: Plaza  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/01624  
; FILING DATE: June 25, 1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A.  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 719 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-01624-31

Query Match	95.0%	Score 3584;	DB 5;	Length 719;
Best Local Similarity	95.3%;	Pred. No. 2.6e-242;		
Matches 704;	Conservative 4;	Mismatches 7;	Indels 24;	Gaps 2;
Qy	1	EVKQENRLLNESSSSOGLLGYFSDLNFOAPVMVTSSTTGDLSIPSELENIPSENQYF	60	
Db	1	EVKQENRLLNESSSSOGLLGYFSDLNFOAPVMVTSSTTGDLSIPSELENIPSENQYF	60	
Qy	61	QSAIWSGFIKYKKSDEYTFATSDADNHVTVMVDQEVINKASNSNKIRLEKGRLLYQIKIY	120	
Db	61	QSAIWSGFIKYKKSDEYTFATSDADNHVTVMVDQEVINKASNSNKIRLEKGRLLYQIKIY	120	
Qy	121	QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPELKQKSNS- - - -RKRSSTAGTVP	176	
Db	121	QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPELKQKSNSNTATIMMQRGNFLQGTVP	180	
Qy	177	DRDNDGIPDSLEVEGYTVDVKNKRTFTFUSPWSINIHKKGLTKYKKSBEKWSASTADPVSDF	236	
Db	181	DRDNDGIPDSLEVEGYTVDVKNKRTFTFUSPWSINIHKKGLTKYKKSBEKWSASTADPVSDF	240	



QY 237 EKVTRIDKXVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTSKNTSTS 296  
DB 241 EKVTRIDKXVSPARHPLVAAYPIVHVDMENIILSKNEDQSTONTDSETRTSKNTSTS 300  
QY 297 RTHTSEVHGNAEVAHSPFDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETWGLNTAD 356  
DB 301 RTHTSEVHGNAEVAHSPFDIGGSVSAFNSNSTVAIDHSLSLAGERTWAETWGLNTAD 360  
QY 357 TARLNANIRYVNTGTAPIYVLPITSLVLGKNOQLATIKADENOLSOILAPNNYPPSKNL 416  
DB 361 TARLNANIRYVNTGTAPIYVLPITSLVLGKNOQLATIKAKENOLSOILAPNNYPPSKNL 420  
QY 417 APIALNAQKQSSPTITMNVNQFLEKTKOLRLDTPQVGNIAATYFNGRVRVDTGSN 476  
DB 421 APIALNAQKQSSPTITMNVNQFLEKTKOLRLDTPQVGNIAATYFNGRVRVDTGSN 460  
QY 477 WSEVLPOIQTETARTIIFNGKDLNVERIAAVNPSDPLETTKPDMTLKEALKIAGFNEP 536  
DB 461 WSEVLPOIQTETARTIIFNGKDLNVERIAAVNPSDPLETTKPDMTLKEALKIAGFNEP 520  
QY 537 NGNLQYQKQDITEFDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNLILIRKRF 596  
DB 521 NGNLQYQKQDITEFDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNLILIRKRF 580  
QY 597 HYDRNNAIVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVI 656  
DB 581 HYDRNNAIVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVI 640  
QY 657 NDYRQDMLNSSLRQDKTFIDFKYNDKPLIYISPNYKVNVAVTKNTIINPSENGDT 716  
DB 641 NDYRQDMLNSSLRQDKTFIDFKYNDKPLIYISPNYKVNVAVTKNTIINPSENGDT 700  
QY 717 STNGIKKILFSSKGYEIG 735  
DB 701 STNGIKKILFSSKGYEIG 719

RESULT 9  
US-09-273-839A-8  
; Sequence 8, Application US/09273839A  
; Patent No. 6329156  
; GENERAL INFORMATION:  
; APPLICANT: Cirino, Nick M  
; APPLICANT: Jackson, Paul J  
; APPLICANT: Lehnert, Bruce E  
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface  
; TITLE OF INVENTION: Receptors  
; FILE REFERENCE: S-89,662  
; CURRENT APPLICATION NUMBER: US/09/273,839A  
; CURRENT FILING DATE: 1999-03-22  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 288  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-273-839A-8  
Query Match 33.8%; Score 1275; DB 4; Length 288;  
Best Local Similarity 100.0%; Pred. No. 1.7e-81; Mismatches 0; Indels 0; Gaps 0;  
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 486 ETTARIIFNGKDLNVERIAAVNPSDPLETTKPDMTLKEALKIAGFNEPNGNLQYQK 545  
DB 29 ETTARIIFNGKDLNVERIAAVNPSDPLETTKPDMTLKEALKIAGFNEPNGNLQYQK 88  
QY 546 DITEFDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNLILIRKRFHYDRNNTAV 605  
DB 89 DITEFDFNFDQTSQNIKNQLAELNATNIYTVLDKIKLNKKNLILIRKRFHYDRNNTAV 148  
QY 606 GADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRYDMLNI 665  
DB 149 GADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIVEIDTEGLKEVINDRYDMLNI 208

QY 666 SSLRODQKTFIDFKYNDKPLIYISPNYKVNVAVTKNTIINPSENGDTSTNGIKKIL 725  
DB 209 SSLRODQKTFIDFKYNDKPLIYISPNYKVNVAVTKNTIINPSENGDTSTNGIKKIL 268  
QY 726 IFSKKGYEIG 735  
DB 269 IFSKKGYEIG 278  
RESULT 10  
US-08-960-780-32  
; Sequence 32, Application US/08960780  
; Patent No. 6204435  
; GENERAL INFORMATION:  
; APPLICANT: Feltelson, Gerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schmeltz, James  
; APPLICANT: Loewer, David  
; APPLICANT: Dullum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide  
; TITLE OF INVENTION: Sequences Which Encode These Toxins  
; NUMBER OF SEQUENCES: 134  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606-6669  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/960,780  
; FILING DATE: 30-OCT-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,848  
; FILING DATE: 30-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: MA-708  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-375-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 32:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 881 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: 177C8  
US-08-960-780-32  
Query Match 20.5%; Score 775; DB 3; Length 881;  
Best Local Similarity 30.2%; Pred. No. 7.2e-46;  
Matches 244; Conservative 134; Mismatches 299; Indels 130; Gaps 34;  
QY 4 QENRLLNESSSQGLLGYFSDLNFCAPMWTSSITGDLISFSSLEN--IFSENCYFQ 61  
DB 39 QKQK---QKMDRKLGLGYFKGKDF-SNLTMPFTRDSTLIYDQQTANKLKKQOEYQ 94  
QY 62 SAWSGFIKVKSDVETFATSDNHTVMWDDQEVINKASNSNKIRLEKRLYQIKIYQ 121

95 SIRWIGLIQSKETGDFTNLSEDEQAIEINGKIISNKGKQVHVLEKGLVPKIEYQ 154  
122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNLOLPELKOKSS-----N 162  
155 SD--TKFNIDSKTPEKELFKIDSONQPOVQOQDELNPFNFNKKESQEFLLAKSKINLFT 212  
163 SRKRSTSGPTVDRNDGIPDSLEVEGYTVVYKKNRTFLSPWISNIHEKGLTKYKSS 222  
213 QKMKREIDED--TDTGDSIPDLWEENGVTI-----QNRIAVKWDDSL-ASKGYTKFVSN 264  
223 PEKSTASDPYSDPEKVTGRDKNVSPKAPLVAAPVIVHVDMENTILSKNEDQSTQNT 282  
265 PLESHTVGDPDYDEKAARDLJLSNAKETFNPLVAAPFVSVNMEKVLSPNENLS----- 320  
283 DSETRTISKNTSTSRTHITSEVHGNAEVSASFDDIGGSVSAGFSNSNS--TVAIDHSL 339  
321 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 363  
340 LAGERTWAETWG-----LNTADTARLANIRVYVGTAPIYVLTPTSLVLGKQTLATIK 395  
364 VAQE--WGTSGNTSQFNTASAGYLNANVRVNYVGTGAIVDVKPTFSFVL--NNDTIAIT 420  
396 ADENQLSOILAPNNYPSKNLAPALNAQKDFSSPTITMNTNQFLEKTKQLRLDQV 455  
421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHSHPTLNKKQVDNLLNNKPMLETNQT 480  
456 YGNIAYNFNGRVRVDTGWSWSEVLPOIQTETARIIFNGKDLNLVERRAAVNPSPLE 515  
481 DG---VYKIKTHGNIVTGGWNGVIOQIKAKTASIIIVDDGE-RVAEKRAAKDYENPED 536  
516 TTKPDMTLKEALKIAF--GFNEPNNGNLOVQKQDITEFDF--NFDQOTSQNIKNQLABL-- 569  
537 KT-PSLTLKDALKLYSPDEIEGLLYKKNPIYESSVMYLDENTAKETVKQLNDTGT 595  
570 ---NATNIYVUDTKLNAKONILIRDRKFHYDRNNAIAGDAESVVKAEHREVINSSTEG 626  
596 KFKDVSHLYDV---KLTPKQNVITIK-LSILYDN---AESNDNSIGKWTNTNIVSGGNG 647  
627 -----LLNID-----KDIRKILSGYIVEIEDTE-----GLKE 654  
648 KQOYSSNPDANLTLNTDAQELKNRDYIISLYMKSEKNTQCEITIDGEIYPTTKTVN 707  
655 VINDRYDMLNI--SSLROGKTFIDFKYKNDKPLIYINPNVKNVYAVTENTINPSE 712  
708 VNKNYKRLDIITAHNIKNPISSIHKT-NDEITLFWDDIST-TDVASIKPEN--LTDSE 763  
713 NGDT-STNGIK---KLIFSKKGYEIG 735  
764 IKQIVSRYGKLEDDGILIDKKGHIHV 790

RESULT 11  
US-09-073-898-32  
Sequence 32, Application US/09073898  
Patent No. 6242669  
GENERAL INFORMATION:  
APPLICANT: Feitelson, Jerald S.  
APPLICANT: Schrepf, H. Ernest  
APPLICANT: Narva, Kenneth E.  
APPLICANT: Stockhoff, Brian A.  
APPLICANT: Schmeits, James  
APPLICANT: Lowey, David  
APPLICANT: Dullum, Charles Joseph  
APPLICANT: Muller-Cohn, Judy  
APPLICANT: Stamp, Lisa  
APPLICANT: Morril, George  
APPLICANT: Finstad-Lee, Stacey  
TITLE OF INVENTION: No. 6242669el Pesticidal Toxins and Nucleotide  
TITLE OF INVENTION: Sequences Which Encode These Toxins  
NUMBER OF SEQUENCES: 144  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606-6669  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/073,898  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/029,848  
FILING DATE: 30-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/960,780  
FILING DATE: 30-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sanders, Jay M.  
REGISTRATION NUMBER: 39,355  
REFERENCE/DOCKET NUMBER: MA-708C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 881 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: PS177C8  
US-09-073-898-32

Query Match 20.5%; Score 775; DB 3; Length 881;  
Best Local Similarity 30.2%; Pred. No. 7,2e-46;  
Matches 244; Conservative 134; Mismatches 299; Indels 130; Gaps 34;

QY 4 QENRLNSESQGLLYVFDLNFQAPMWTSTTGLSIPSSSELEN--IPSENQYEQ 61  
DB 39 QKQK---QKMDRKGULGYFKGKDF-SNLTFAPTRDSTLIYDQOTANKLLDKQEQVQ 94  
QY 62 SAWSGFIYKKSDEYTFATSNADNHVMTVMDQEVINKASNKIRLEKRLYQIKIQV 121  
DB 95 SIRWIGLIQSKETGDFTNLSEDEQAIEINGKIISNKGKQVHVLEKGLVPKIEYQ 154  
QY 122 RENPTEKGLD-----FKLYWTDSONKKEVSSDNLOLPELKOKSS-----N 162  
DB 155 SD--TKFNIDSKTPEKELFKIDSONQPOVQOQDELNPFNFNKKESQEFLLAKSKINLFT 212  
QY 163 SRKRSTSGPTVDRNDGIPDSLEVEGYTVVYKKNRTFLSPWISNIHEKGLTKYKSS 222  
DB 213 QKMKREIDED--TDTGDSIPDLWEENGVTI-----QNRIAVKWDDSL-ASKGYTKFVSN 264  
QY 223 PEKSTASDPYSDPEKVTGRDKNVSPKAPLVAAPVIVHVDMENTILSKNEDQSTQNT 282  
DB 265 PLESHTVGDPDYDEKAARDLJLSNAKETFNPLVAAPFVSVNMEKVLSPNENLS----- 320  
QY 283 DSETRTISKNTSTSRTHITSEVHGNAEVSASFDDIGGSVSAGFSNSNS--TVAIDHSL 339  
DB 321 ---NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHS 363  
QY 340 LAGERTWAETWG-----LNTADTARLANIRVYVGTAPIYVLTPTSLVLGKQTLATIK 395  
DB 364 VAQE--WGTSGNTSQFNTASAGYLNANVRVNYVGTGAIVDVKPTFSFVL--NNDTIAIT 420  
QY 396 ADENQLSOILAPNNYPSKNLAPALNAQKDFSSPTITMNTNQFLEKTKQLRLDQV 455  
DB 421 AKSNSTALNISPGESYPKKGQNGIAITSMDDFNHSHPTLNKKQVDNLLNNKPMLETNQT 480

456 YGNIATYVFNENGRVVRVDTGSNWSEVLPQIQTETARIIFNGKLNLVERRIAAVNPDPLE 515  
481 DG---VYKIDTHGNIVTGEWNGVIOQIKAKTASIIVDGGE-RVAEKRAAKDYENPED 536  
516 TTKPDMTLKALKIAF--GFNEPNGNLOYQKIDITEPFD--NFDQOOTSQNIKNQLAEL-- 569  
537 KT-PSLTLDKALKUSYDPEKEIEGLLYYKPKPIYESSVMTYLDENTAKEVTQLNDTTG 595  
570 ---NATNIYTVLDKIKLNKAKNLIIRDKRPHYDRNNIAVGADSVVKEAHEVINSSTEG 626  
596 KFKDVSHLYDV---KLTPKGNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGG 647  
627 -----LLNID-----KDIRKILSGYIVEIDTE-----GLKE 654  
648 KQYSSNNPDANLTNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGIBYPTTKVN 707  
655 VINDRYMLNI--SSLRODQKTIIDFKYNDKPLIYSNPNYKVVAVTKENTINPSE 712  
708 VNKDNYKRLDIIAHNIKSNPISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763  
713 NGDT-STNGIK---KILIFSKGYEIG 735  
764 IKQIYSRYGKLEPGLIDKGGIHYG 790

RESULT 12  
US-09-307-106-8  
; Sequence 8, Application US/09307106  
; Patent No. 6603063  
; GENERAL INFORMATION:  
; APPLICANT: Feitelson, Jerald S.  
; APPLICANT: Schnepf, H. Ernest  
; APPLICANT: Narva, Kenneth E.  
; APPLICANT: Stockhoff, Brian A.  
; APPLICANT: Schweits, James  
; APPLICANT: Loewer, David  
; APPLICANT: Dullum, Charles Joseph  
; APPLICANT: Muller-Cohn, Judy  
; APPLICANT: Stamp, Lisa  
; APPLICANT: Morrill, George  
; APPLICANT: Finstad-Lee, Stacey  
; TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide  
; TITLE OF INVENTION: Sequences Which Encode These Toxins  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606-6669  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/09/307,106  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/029,848  
; FILING DATE: 30-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/960,780  
; FILING DATE: 30-OCT-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/073,898  
; FILING DATE: 05-MAY-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sanders, Jay M.  
; REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 881 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: PSI177C8a  
US-09-307-106-8

Query Match 20.5%; Score 775; DB 4; Length 881;  
Best Local Similarity 30.2%; Pred. No. 7, 2e-46;  
Matches 244; Conservative 134; Mismatches 299; Indels 130; Gaps 34;  
4 QENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELEN--IPSENQVFO 61  
39 QKNQ---QKEMDRKGLLYYFKGDF-SNLTAFAPTRDSTLIYDQQTANKLDDKKQOEYQ 94  
62 SAIWSGFIKVKKSDEYTFATSDAHNVHMVDDQEVINKASNSKIRLEKGRLYQIKIQYQ 121  
95 SIRWIGLIOSKETGDTFNLSEDEQAIIIEINGKIIISNKGKEQVHVLEKGLVPIKIEYQ 154  
122 RENPTEKGLD-----FKLWYTDSONKKEVISSDNLQLPKQKSS-----N 162  
155 SD--TKFNIDSKTFKELKLFKIDSONQPOQVQDDELNPENFKESQBFLLAKPSKINLFT 212  
163 SRKGRSTAGTPVDRDNDGIDPSLEVEGYTVVKNKRTFLSPWISNHEKGLTKYKSS 222  
213 QKMKREIDED---TDTGDSIPDLWEENGYTI---QNRIVAKWDDSL-ASKGYTKFVSN 264  
223 PEKMWSTADSPYDFEKVTGRIDKNSPEARHPLVAAYPIVHVDMENIILSKNEDOSTQMT 282  
265 PLESHTVGDPYTDYKAAARDLDSNAKETFELVAAPPSVNVSMKVILSPNENLS---- 320  
283 DSETRIISKNTSRTHTSEVGNAAEVHASFDTIGGSYSAGFSNSNS---TVAIDHSLS 339  
321 ---NSVESHSSTNWSYT-----NTB-----GASVEAGIGPKGISFGVSUYNQHSFT 363  
340 LAGERTWAETMG---LNTADTARLNANIRYVNTGTAPIYNNVLTPTTSLVLGKXQTLATIK 395  
364 VAQE--WGTSTGNTSQFNTASAGLVANVRNNVGTGAIVDVKPTTSFVL--NNDTIATIT 420  
396 ADENQLSQILAPNNYTPSKNLAIPALNAOKDPSSTFITWYNNQFLEKTKQLRLDQV 455  
421 AKSNSTALNISPGESYFKGQNGIAITSMDDFNHPITLKKQVDNLLNNKPMLETNQT 480  
456 YGNIATYVFNENGRVVRVDTGSNWSEVLPQIQTETARIIFNGKLNLVERRIAAVNPDPLE 515  
481 DG---VYKIDTHGNIVTGEWNGVIOQIKAKTASIIVDGGE-RVAEKRAAKDYENPED 536  
516 TTKPDMTLKALKIAF--GFNEPNGNLOYQKIDITEPFD--NFDQOOTSQNIKNQLAEL-- 569  
537 KT-PSLTLDKALKUSYDPEKEIEGLLYYKPKPIYESSVMTYLDENTAKEVTQLNDTTG 595  
570 ---NATNIYTVLDKIKLNKAKNLIIRDKRPHYDRNNIAVGADSVVKEAHEVINSSTEG 626  
596 KFKDVSHLYDV---KLTPKGNVTIK-LSILYDN---AESNDNSIGKWTNTNIVSGNGG 647  
627 -----LLNID-----KDIRKILSGYIVEIDTE-----GLKE 654  
648 KQYSSNNPDANLTNTDAQEKLNKRDYIISLYMKSEKNTQCEITIDGIBYPTTKVN 707  
655 VINDRYMLNI--SSLRODQKTIIDFKYNDKPLIYSNPNYKVVAVTKENTINPSE 712  
708 VNKDNYKRLDIIAHNIKSNPISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763  
713 NGDT-STNGIK---KILIFSKGYEIG 735

Db 764 IKQYRYGKLEBGLIDKGGIHYG 790

RESULT 13

US-09-850-351A-32

Sequence 32, Application US/09850351A

Patent No. 6656908

GENERAL INFORMATION:

APPLICANT: Feitelson, Jerald S.

Schneff, H. Ernest

Narva, Kenneth E.

Stockhoff, Brian A.

Schmeits, James

Loewer, David

Dullum, Charles Joseph

Muller-Cohn, Judy

Stamp, Lisa

Morrill, George

TITLE OF INVENTION: No. 6656908el Pesticidal Toxins and Nucleotide Sequences Which Encode These Toxins

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606-6669

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/850,351A

FILING DATE: 07-May-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/073,898

FILING DATE: 06-MAY-1998

APPLICATION NUMBER: US 08/960,780

FILING DATE: 30-OCT-1997

APPLICATION NUMBER: US 60/029,848

FILING DATE: 30-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708CD1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 881 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: PSI77C8

SEQUENCE DESCRIPTION: SEQ ID NO: 32:

US-09-850-351A-32

Query Match 20.5%; Score 775; DB 4; Length 881;

Best Local Similarity 30.2%; Pred. No. 7.2e-46;

Matches 244; Conservative 13; Mismatches 299; Indels 130; Gaps 34;

Qy 4 QENRLNESSSQGLGYFSDLNFOAPMVVTSSTGDSLSPSSLEN--IPSENQYFQ 61

Db 39 QKQKQ--QKEMDRKGLLGYYFKGKDF-SNLTMEAPTRDSPLIYDQQTANKLIDKQKQYQ 94

Qy 62 SATWSGFIKVKSDSEYFATSAADNHTVWDDQDEVINKASNSKIRLEKRLVQIKIQY 121

Db 95 SIRWIGLIQSKETGDTFENLSEDEQAITEINGKIIISNKGKEQVVHLEKGLKLVPIKIEYQ 154

Qy 122 RENPTEKGLD-----FKLYWTDSONKEVSSDNLQLPKQKSS-----N 162

Db 155 SD--TKFNIDSKTFKELKFKIDSONQFQQVQOQDELNPENFKESQBFAPKPSKINLFT 212

Qy 163 SRKKESTAGPTVPDRDNDGIPDSLEVEGYTVDVKNKETFSLSPWISNHEKKGITKYS 222

Db 213 QRMKEIDED--TDTGDSIPDLWEENGYTI---QNRIAVKWDDSL-ASKGYTKFYSN 264

Qy 223 PEKWTASDPYSDPEKVTGRIDKNVSPARHPLVAAPYIVHVDMBENIILSKNEQDQSTQNT 282

Db 265 PLESHTVGDPYTDYEKAARDLDSNAKETFPNLAAPFVSNVSMKVLSPNENLUS---- 320

Qy 283 DSETRTISKNTSTSTRTHTSEVHGNAEVAHSPFDIGGSVAGFSNSNSS---TVAIDHLS 339

Db 321 ----NSVESHSSTNWSYT-----NTE-----GASVEAGIGPKGISFGVSVNYQHSET 363

Qy 340 LAGERTWARTMG----LNTADTARLANANIRYNTGTAPIYVLPITTSVLGKNGQTLATIK 395

Db 364 VAQE--WGTSTGNTSQFNTASAGYLNANRVYNNVGTGAIYDVKPTISFVL--NNDTIAIT 420

Qy 396 ADENQLSQILAPNNYVPSKNLAPIALNAQKSPSTPTIMYNNQFLELEKTKQLRLDQV 455

Db 421 AKSNSTALNISGESYPKGGQNGIAITSDMDFNSHPITLNKKQVDNLLNNKPMLETNQT 480

Qy 456 YGNIATYFNGRVRVDTGNSWSEVLPOIQTETARIIIFNGKDLNLVERIAAVNPSDPLE 515

Db 481 DG---VYKIKDTHGNIVTGEWNGVITQIKAKTASIIIVDDGE-RVAEKRVAAKDYENPED 536

Qy 516 TTKPDMTLKALKIAP--GFNEPNGNLOYQKDIETEDF--NFDQOTSONIKNQLAEL-- 569

Db 537 KT-PSUTLKDALKLSYDPDEIKIEGLLYYKNPYESSVMTYLDENTAKEVTKQLNDITG 595

Qy 570 ---NATNIYTVLDKIKLNAMNILIRDKRFHVRNNAIAGADESVVKEAHRVINSSTEG 626

Db 596 KFKDVSHLYDV---KLTPEKMNVTIK-LSILYDN---AESNDNSIGKNTNTNIVSGNNG 647

Qy 627 -----LLNLID-----KDIRKILSYIVEIEDTE-----GLKE 654

Db 648 KKOYSSNNPDANLTNTDAQEKLNRDYYISLYMKSEKNTQCEITIDGEIYPIITTKTVN 707

Qy 655 VINDRYDMLNI--SSLRQDQKTFIDPKYNDKLPYISNPNYKVNKYATKENTINPSE 712

Db 708 VNKONYKRLDIIAHTIKSNPISIIHKT-NDEITLFWDDISI-TDVASIKPEN--LTDSE 763

Qy 713 NGDT-STNGIK---KILIFSKKGVEIG 735

Db 764 IKQYRYGKLEBGLIDKGGIHYG 790

RESULT 14

US-08-471-033-5

Sequence 5, Application US/08471033

Patent No. 5770596

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W

APPLICANT: Kozziel, Michael G

APPLICANT: Mullins, Martha A

APPLICANT: Nye, Gordon J

APPLICANT: Carr, Brian

APPLICANT: Dessai, Nalini M

APPLICANT: Kostichka, N. Kristy

APPLICANT: Duck, Nicholas B

APPLICANT: Estruch, Juan J

TITLE OF INVENTION: No. 5770596el Pesticidal Proteins and Strains

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: CIBA-GEIGY Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA



Search completed: May 3, 2004, 19:42:42  
Job time : 17.0986 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:40:13 ; Search time 36.6393 Seconds  
(without alignments)  
5560.545 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKKILIFSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277199581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3761	99.7	735	15	US-10-410-647-30
2	3761	99.7	735	12	US-09-848-909-1
3	3761	99.7	735	12	US-09-848-909-2
4	3761	99.7	735	12	US-09-848-909-3
5	3761	99.7	735	12	US-09-848-909-4
6	3761	99.7	735	12	US-09-848-909-5
7	3761	99.7	735	12	US-09-848-909-6
8	3761	99.7	735	12	US-09-848-909-7
9	3761	99.7	735	12	US-09-848-909-8
10	3761	99.7	735	12	US-09-848-909-9
11	3761	99.7	735	12	US-09-848-909-10
12	3761	99.7	735	12	US-09-848-909-11
13	3761	99.7	735	12	US-09-848-909-12
14	3761	99.7	735	12	US-09-848-909-13
15	3761	99.7	735	12	US-09-848-909-14

16	3761	99.7	736	12	US-09-848-909-15	Sequence 15, Appl
17	3761	99.7	736	12	US-09-848-909-16	Sequence 16, Appl
18	3761	99.7	736	12	US-09-848-909-17	Sequence 17, Appl
19	3761	99.7	736	12	US-09-848-909-18	Sequence 18, Appl
20	3761	99.7	736	12	US-09-848-909-19	Sequence 19, Appl
21	3761	99.7	736	12	US-09-848-909-20	Sequence 20, Appl
22	3761	99.7	736	12	US-09-848-909-21	Sequence 21, Appl
23	3761	99.7	736	12	US-09-848-909-22	Sequence 22, Appl
24	3761	99.7	736	15	US-10-442-502-7	Sequence 7, Appl
25	3761	99.7	763	15	US-10-442-502-5	Sequence 5, Appl
26	3761	99.7	764	15	US-10-442-502-6	Sequence 6, Appl
27	3758	99.6	735	12	US-10-402-466A-9	Sequence 9, Appl
28	3758	99.6	735	12	US-10-402-466A-13	Sequence 13, Appl
29	3758	99.6	735	14	US-10-332-282-13	Sequence 13, Appl
30	3754	99.5	735	12	US-09-848-909-30	Sequence 30, Appl
31	3754	99.5	764	12	US-10-253-286-681	Sequence 681, App
32	3754	99.5	764	15	US-10-245-871-681	Sequence 681, App
33	3654	96.8	764	9	US-09-747-521-4	Sequence 4, Appl
34	3654	96.8	764	13	US-10-106-014-4	Sequence 4, Appl
35	3654	96.8	764	13	US-10-105-695-4	Sequence 4, Appl
36	3654	96.8	764	14	US-10-105-694-4	Sequence 4, Appl
37	3050	80.8	599	12	US-09-848-909-24	Sequence 24, Appl
38	3045	80.7	595	14	US-10-332-282-11	Sequence 11, Appl
39	2918	77.3	573	12	US-10-402-466A-22	Sequence 22, Appl
40	2906	77.0	569	15	US-10-442-502-8	Sequence 8, Appl
41	2496	66.1	487	14	US-10-332-282-5	Sequence 5, Appl
42	2181	57.8	426	14	US-10-332-282-9	Sequence 9, Appl
43	2135	56.6	423	12	US-10-402-466A-24	Sequence 24, Appl
44	1632	43.2	318	14	US-10-332-282-7	Sequence 7, Appl
45	1338	35.5	258	14	US-10-332-282-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-410-647-30  
; Sequence 30, Application US/10410647  
; Publication No. US20030235818A1  
; GENERAL INFORMATION:  
; APPLICANT: PLEXUS VACCINE, INC.  
; APPLICANT: Katritch, Vsevolod  
; APPLICANT: Bordner, Andrew  
; APPLICANT: Deans, Robert  
; APPLICANT: Sumner, Mary  
; TITLE OF INVENTION: IMMUNOGENIC PEPTIDES, AND METHOD OF IDENTIFYING SAME  
; FILE REFERENCE: PLEX1110-1  
; CURRENT APPLICATION NUMBER: US/10/410,647  
; CURRENT FILING DATE: 2003-04-08  
; PRIOR APPLICATION NUMBER: US 60/373,668  
; PRIOR FILING DATE: 2002-04-17  
; PRIOR APPLICATION NUMBER: US 60/371,256  
; PRIOR FILING DATE: 2002-04-08  
; PRIOR APPLICATION NUMBER: US 60/371,250  
; PRIOR FILING DATE: 2002-04-08  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 30  
; TYPE: PRT  
; LENGTH: 735  
; ORGANISM: Bacillus anthracis  
US-10-410-647-30

Query Match 99.7%; Score 3761; DB 15; Length 735;  
Best Local Similarity 99.7%; Pred. No. 4.3e-272;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENLLNESSSQGLLYVFDLNFQAPMVVTSSTGDLSTPSSSELENPSENQV 60

Db 1 EVKQENLLNESSSQGLLYVFDLNFQAPMVVTSSTGDLSTPSSSELENPSENQV 60

QY 61 QSAWISGFIKVKKSDEYTFATSDADNHTVMVDQEVINKASNKIRLEKGLYQIKQY 120

Db 61 QSAIWSGFIKVKSSDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLQPLKOKSSNSRKRSTSGPTVPDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLQPLKOKSSNSRKRSTSGPTVPDRDN 180  
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDYSDFEKT 240  
Db 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEYHAGFFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTMTGLTADTARL 360  
Db 301 SEVHGNAEYHAGFFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTMTGLTADTARL 360  
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVVDYTGNSWSEV 480  
Db 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVVDYTGNSWSEV 480  
QY 481 LPOIQETTARIIFNGKDLNVERRTAAVNPSPDLETTKPDMTLKEALKIAGFNEPKNL 540  
Db 481 LPOIQETTARIIFNGKDLNVERRTAAVNPSPDLETTKPDMTLKEALKIAGFNEPKNL 540  
QY 541 QYQKDIETEFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600  
Db 541 QYQKDIETEFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 2

US-09-848-909-1

; Sequence 1, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/060002 and Prevention of Bacterial Infection

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-1

Query Match

Best Local Similarity 99.7%; Score 3761; DB 12; Length 736;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNESESSQGLLYFFDLNFOAPMVVTSSTGDLSPSELENIPSENYVF 60  
Db 1 EVKQENRLNESESSQGLLYFFDLNFOAPMVVTSSTGDLSPSELENIPSENYVF 60  
QY 61 QSAIWSGFIKVKSSDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
Db 61 QSAIWSGFIKVKSSDEYTFATSDADNHVTMWDDQEVINKASNSKIRLEKGRLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLQPLKOKSSNSRKRSTSGPTVPDRDN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVYSSDNQLQPLKOKSSNSRKRSTSGPTVPDRDN 180  
QY 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDYSDFEKT 240  
Db 181 DGPDSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
QY 301 SEVHGNAEYHAGFFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTMTGLTADTARL 360  
Db 301 SEVHGNAEYHAGFFDIGGSVAGFSNSSTVAIDHSLSLAGERTWAGTMTGLTADTARL 360  
QY 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVVDYTGNSWSEV 480  
Db 421 LNAQKDFSTPTTMNYNQFLEKTKQLRLDQDVYGNATYFNENGRVVDYTGNSWSEV 480  
QY 481 LPOIQETTARIIFNGKDLNVERRTAAVNPSPDLETTKPDMTLKEALKIAGFNEPKNL 540  
Db 481 LPOIQETTARIIFNGKDLNVERRTAAVNPSPDLETTKPDMTLKEALKIAGFNEPKNL 540  
QY 541 QYQKDIETEFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600  
Db 541 QYQKDIETEFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKLAKMILIRDKRFHYDR 600  
QY 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAAREVINSSTGLLNLNDKDIRKILSGYVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNSSLRQDGKTFIDFKKNDKPLVYISNPNYKVVYAVTKENTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 3

US-09-848-909-2

; Sequence 2, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/060002 and Prevention of Bacterial Infection

; CURRENT APPLICATION NUMBER: US/09/848,909

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis



```
US-09-848-909-2
Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLLNESSESSQGLGYFSDNLFQAPMVVTSITGDLSPSSSELENIPSENOYF 60
DB 1 EVKQENRLLNESSESSQGLGYFSDNLFQAPMVVTSITGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKSDYTFATSDAHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYTFATSDAHVMTWDDQEVINKASNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPFLKQKSSNRKKESTASGPTVDPDRN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPFLKQKSSNRKKESTASGPTVDPDRN 180
QY 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240
DB 181 DGIIPDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTASDPYDFEKT 240
QY 241 GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
DB 241 GRIDKNVSPARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
QY 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDDIGSVSAGFSNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSNLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSNLAPIA 420
QY 421 LNAQKDFSSPTIMYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSNLAPIA 480
DB 421 LNAQKDFSSPTIMYVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNPNYPKSNLAPIA 480
QY 481 LPOIQTETARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
DB 481 LPOIQTETARIIFNGKOLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
QY 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDDQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKIRKILSGYIIVEIDTEGLKEVINDRY 660
DB 601 NNIAGADESVVKEAHREVINSSTEGLLNIDKIRKILSGYIIVEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKYNDKPLPLYSNPKNVYAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKYNDKPLPLYSNPKNVYAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

RESULT 4
US-09-848-909-3
; Sequence 3, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04

US-09-848-909-4
; Sequence 4, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
```

```
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-4

Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTSGAGTVPDRDN 180
Db 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTSGAGTVPDRDN 180
QY 181 DGIPLSLEVEGYTVVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYDPEKVT 240
Db 181 DGIPLSLEVEGYTVVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYDPEKVT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
QY 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDVYGNIAATYNFENGRVVRDTSNWSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDVYGNIAATYNFENGRVVRDTSNWSSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
QY 541 QYQKDIITBFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIITBFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHREVNSTEGLLINIDKDIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVNSTEGLLINIDKDIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSNG 720
QY 721 IKKILFSSKKGYEIG 735
Db 721 IKKILFSSKKGYEIG 735
```

RESULT 6

US-09-848-909-5

```
; Sequence 5, Application US/09848909
; Publication No. US20020039598A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-5
```

```
Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
Db 1 EVKQENRLNSESSESSQGLLYFFSDLNFPQAPMVVTSSTTGDLSIPSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFTKVKSDYEYFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTSGAGTVPDRDN 180
Db 121 QRENPTKGLDPKLYWTDSONKKEVISSDNLQLPKQKSSNSRKRSTSGAGTVPDRDN 180
QY 181 DGIPLSLEVEGYTVVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYDPEKVT 240
Db 181 DGIPLSLEVEGYTVVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASPYDPEKVT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
Db 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRTHT 300
QY 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAGTGLNTADTARL 360
QY 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVNTGTAPIYNNVPTTSLVLGKQTLATIKADENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDVYGNIAATYNFENGRVVRDTSNWSSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQDVYGNIAATYNFENGRVVRDTSNWSSEV 480
QY 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENGNL 540
QY 541 QYQKDIITBFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIITBFDNFDOQTSONIKNQLAELNATNIYTVLDKIKNAKNNILIRDKRPHYDR 600
QY 601 NNIAVGADESUVKEAHREVNSTEGLLINIDKDIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVNSTEGLLINIDKDIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSNG 720
Db 661 DMLNISSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAATKENTIIINPSENGDTSNG 720
QY 721 IKKILFSSKKGYEIG 735
```

```
Db 721 IKKILFSSKGYEIG 735
|||||
RESULT 7
US-09-848-909-6
; Sequence 6, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-6

Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLLNESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLLNESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTPVDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTPVDRDN 180
Qy 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240
Db 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPENGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPENGNL 540
Qy 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKEFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKEFHYDR 600
Qy 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKRKILSGVIEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKRKILSGVIEIEDTEGLKEVINDRY 660
Qy 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKRKILSGVIEIEDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDKRKILSGVIEIEDTEGLKEVINDRY 660
```

```
Qy 661 DMLNISLRODGKTFIDFKKYNKPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISLRODGKTFIDFKKYNKPLYISNPNYKVVYAVTKENTIIINPSENGDTSTNG 720
Qy 721 IKKILFSSKGYEIG 735
Db 721 IKKILFSSKGYEIG 735

RESULT 8
US-09-848-909-7
; Sequence 7, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-7

Query Match          99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLLNESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Db 1 EVKQENRLLNESSESSQGLLYGFFDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
Qy 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKSDDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTPVDRDN 180
Db 121 QRENPTKEGLDFKLYWTDSONKKEVSSDNLQLPKQKSSNSRKRKSTSGAGTPVDRDN 180
Qy 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240
Db 181 DGIPTDSLEVEGYTVVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASTADPYSDPEKVT 240
Qy 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Db 241 GRIDKNVSPPEARHPLVAAVPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300
Qy 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHAGFFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVLGKNOTLATIKADENQLSOILAPNNYPSKNLAPIA 420
Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNITATYNFENGRVRVDTGSNWSEV 480
Qy 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPENGNL 540
Db 481 LPOIQETTARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAPFNEPENGNL 540
Qy 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKEFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLNAMNIIIRDKEFHYDR 600
```

```
Db 541 QYQKLTDFDFNFDQTSQNKQLAELNATNITVLDKIKLNAKNNILIRDRFHYDR 600
Qy 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 9
US-09-848-909-8
; Sequence 8, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-8

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPESENYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPESENYF 60
Qy 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKOKSSNRKRSSTAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKOKSSNRKRSSTAGTPVDRDN 180
Qy 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
Qy 421 LNAQKQFSSTPTIMYNNQFIELEKTKQLRLDITDQVYGNITATYFNENGRVVDTGSNWSEV 480
Db 421 LNAQKQFSSTPTIMYNNQFIELEKTKQLRLDITDQVYGNITATYFNENGRVVDTGSNWSEV 480
```

```
Qy 481 LPQIQTETARIIPNGKOLNVERIAAVNPSPDPLETTKPDWTLKALKIAPGNEPNGL 540
Db 481 LPQIQTETARIIPNGKOLNVERIAAVNPSPDPLETTKPDWTLKALKIAPGNEPNGL 540
Qy 541 QYQKDIITFEFDFNFDQTSQNKQLAELNATNITVLDKIKLNAKNNILIRDRFHYDR 600
Db 541 QYQKDIITFEFDFNFDQTSQNKQLAELNATNITVLDKIKLNAKNNILIRDRFHYDR 600
Qy 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAVGADESUVKEAHEVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660
Qy 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
Db 661 DMLNSSLRQDGKTFIDFKKYNKPLYSNPNYKVNVAVTKENTIIINPSENGDSTNG 720
Qy 721 IKKILIFSKGYEIG 735
Db 721 IKKILIFSKGYEIG 735

RESULT 10
US-09-848-909-9
; Sequence 9, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909-9

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPESENYF 60
Db 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSELENIPESENYF 60
Qy 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIVKVKSDDEYTFATSDADNHVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKOKSSNRKRSSTAGTPVDRDN 180
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKOKSSNRKRSSTAGTPVDRDN 180
Qy 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240
Db 181 DGIPODSEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSSTASDPYDFEKT 240
Qy 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRHT 300
Db 241 GRIDKNVSPPEARHPLVAAPYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSSTRHT 300
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETWGLNTADTARL 360
Qy 361 NANIRVYNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
```

```
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
QY 481 LPOIETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPOIETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGNEPENG 540
QY 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 11
US-09-848-909-10
; Sequence 10, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 736
; TYPE: PR1
; ORGANISM: Bacillus anthracis
US-09-848-909-10

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLOLPELKOKSSNSRKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLOLPELKOKSSNSRKRSTAGPTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
QY 241 GRIDKNVSPEARHPLVAAPVIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSTRHT 300
Db 241 GRIDKNVSPEARHPLVAAPVIVHVDMENIILSKNEDQSTQNTDSTRITSKNTSTRHT 300
```

```
QY 301 SEVHGNAEVHASFDFIGGSVSAFSSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDFIGGSVSAFSSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNOTLATIKAKENQLSQILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYNFENGVRVDTGSNWSEV 480
QY 481 LPOIETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGNEPENG 540
Db 481 LPOIETTARIIFNGKDLNVERRIAAVNPSPLETTKPDMTLKEALKIAFGNEPENG 540
QY 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
Db 541 QYQKDIETEDFNFDOOTSQNIKNQLAELNATNIYVLDKIKLNAKNNILIRDKRPHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
QY 721 IKKILIFSKKGYEIG 735
Db 721 IKKILIFSKKGYEIG 735

RESULT 12
US-09-848-909-11
; Sequence 11, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 736
; TYPE: PR1
; ORGANISM: Bacillus anthracis
US-09-848-909-11

Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTGDLSPSSSELENIPSENOYF 60
QY 61 QSAIWSGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDYTFATSNADNHVTMWVDDQEVINKASNNKIRLEKGLYQIKIY 120
QY 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLOLPELKOKSSNSRKRSTAGPTVPDRDN 180
Db 121 QRENTEKGLDFKLYWTDSONKEVIVSSDNLOLPELKOKSSNSRKRSTAGPTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTDPYDFEKT 240
```

Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKASTASDPYSDFEKT 240  
Qy 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Qy 301 SEVHGNAEVAHAFDILGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHAFDILGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRTGNSWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRTGNSWSEV 480  
Qy 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
Qy 541 QYQKDIITFDFNFDOOTSONKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600  
Db 541 QYQKDIITFDFNFDOOTSONKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600  
Qy 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 13

US-09-848-909-12

; Sequence 12, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-12

Query Match

Best Local Similarity 99.7%; Score 3761; DB 12; Length 736;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFFDLNFAQPMVVTSTTGDLSIPSELENIIPSENQYF 60  
Db 1 EVKQENRLNSESQGLGYFFDLNFAQPMVVTSTTGDLSIPSELENIIPSENQYF 60  
Qy 61 QSAIWSGFIKVKKSDEYFATSDADNHVTWVDDQEVINKANSNKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYFATSDADNHVTWVDDQEVINKANSNKIRLEKGLYQIKIY 120

Qy 121 QRENTEKGLDFKLYNTDSQNKKEVISSDNLQLPKLKQSSNRKRSAGTSAGTVPDRDN 180  
Db 121 QRENTEKGLDFKLYNTDSQNKKEVISSDNLQLPKLKQSSNRKRSAGTSAGTVPDRDN 180  
Qy 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKASTASDPYSDFEKT 240  
Db 181 DGIPTDSLEVEGYTVDVKNKRTFLSPWISNHEKKGLTKYKSSPEKASTASDPYSDFEKT 240  
Qy 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Db 241 GRIDKNVSPPEARPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
Qy 301 SEVHGNAEVAHAFDILGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHAFDILGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Qy 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRTGNSWSEV 480  
Db 421 LNAQKDFSSPTITMNYNQFLELEKTKQLRLDQVYGNATYVNFENGRVVRTGNSWSEV 480  
Qy 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
Db 481 LPOIQETTARIIFNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
Qy 541 QYQKDIITFDFNFDOOTSONKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600  
Db 541 QYQKDIITFDFNFDOOTSONKQLAELNATNIYVLDKIKLNAMNILLIRDKRPHYDR 600  
Qy 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNIAGADESVVKEAAREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Db 661 DMLNISSLRQDGKTFIDFKYNDKPLIYSNPNYKVNVAVTKENTIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 14

US-09-848-909-13

; Sequence 13, Application US/09848909

; Publication No. US20020039588A1

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; APPLICANT: Sellman, Brett R.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909

; PRIOR FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13

; LENGTH: 736

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909-13

Query Match

Best Local Similarity 99.7%; Score 3761; DB 12; Length 736;

Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFFDLNFAQPMVVTSTTGDLSIPSELENIIPSENQYF 60

```
Db 1 EVKQENRLLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHEVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHEVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTASAGTVPDRDN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTASAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYSDFEVT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYSDFEVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYNYQFLELEKTKQLRLDQVYGNATYVFNENGRVVDVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITWYNYQFLELEKTKQLRLDQVYGNATYVFNENGRVVDVDTGSNWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKQALKIAFGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKQALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMNILLRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMNILLRDKRFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

## RESULT 15

```
US-09-848-909-14
; Sequence 14, Application US/09848909
; Publication No. US20020039588A1
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; APPLICANT: Sellman, Brett R.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/06002
; CURRENT APPLICATION NUMBER: US/09/848,909
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Bacillus anthracis
```

US-09-848-909-14

```
Query Match 99.7%; Score 3761; DB 12; Length 736;
Best Local Similarity 99.7%; Pred. No. 4.3e-272;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 EVKQENRLLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
Db 1 EVKQENRLLNESSESSQGLLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHEVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHEVTMWVDDQEVINKASNSNKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDPLKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTASAGTVPDRDN 180
Db 121 QRENPTKGLDPLKLYWTDSONKKEVITSSDNLOLPELKQKSSNSRKRSTASAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYSDFEVT 240
Db 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGITKYKSSPEKWSASTASDPYSDFEVT 240
QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNSTSRHT 300
QY 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
Db 301 SEVHGNAEVHASFDDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
Db 361 NANIRYVNTGTAPIYVNLPTTSLVLGKQTLATIKADENQLSQIILAPNNYPSKNLAPIA 420
QY 421 LNAQKDFSSPTITWYNYQFLELEKTKQLRLDQVYGNATYVFNENGRVVDVDTGSNWSEV 480
Db 421 LNAQKDFSSPTITWYNYQFLELEKTKQLRLDQVYGNATYVFNENGRVVDVDTGSNWSEV 480
QY 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKQALKIAFGFNEPENG 540
Db 481 LPQIQETTARIIFNGKDLNLVERRIAAVNPSPLETTKPDMTLKQALKIAFGFNEPENG 540
QY 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMNILLRDKRFHYDR 600
Db 541 QYQKDIITEFDNFDOQTSONIKNQLAELNATNIYTVLDKIKLNKAMNILLRDKRFHYDR 600
QY 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
Db 601 NNIAGADESVVKEAHEVINSSTEGLLNIDKIRKILSGYIVIEDETEGLKEVINDRY 660
QY 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
Db 661 DMLNISSLRODGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
Db 721 IKKILIFSCKGYEIG 735
```

Search completed: May 3, 2004, 20:02:09  
Job time : 41.6393 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:36:13 ; Search time 175.546 Seconds  
(without alignments)  
4086.665 Million cell updates/sec

Title: US-09-848-909A-10  
Perfect score: 3774  
Sequence: 1 EVKQENLLNESSSQGLL.....TSTNGIKLILFKKGYEIG 735

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA Main:

- 1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pcp.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pcp.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pcp.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pcp.\*
- 5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pcp.\*
- 6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pcp.\*
- 7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pcp.\*
- 8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pcp.\*
- 9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pcp.\*
- 10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pcp.\*
- 11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pcp.\*
- 12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pcp.\*
- 13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pcp.\*
- 14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pcp.\*
- 15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pcp.\*
- 16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pcp.\*
- 17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pcp.\*
- 18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pcp.\*
- 19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pcp.\*
- 20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pcp.\*
- 21: /cgn2\_6/ptodata/2/paa/US097A\_COMB.pcp.\*
- 22: /cgn2\_6/ptodata/2/paa/US097B\_COMB.pcp.\*
- 23: /cgn2\_6/ptodata/2/paa/US098\_COMB.pcp.\*
- 24: /cgn2\_6/ptodata/2/paa/US099A\_COMB.pcp.\*
- 25: /cgn2\_6/ptodata/2/paa/US099B\_COMB.pcp.\*
- 26: /cgn2\_6/ptodata/2/paa/US100\_COMB.pcp.\*
- 27: /cgn2\_6/ptodata/2/paa/US101\_COMB.pcp.\*
- 28: /cgn2\_6/ptodata/2/paa/US102\_COMB.pcp.\*
- 29: /cgn2\_6/ptodata/2/paa/US103\_COMB.pcp.\*
- 30: /cgn2\_6/ptodata/2/paa/US104\_COMB.pcp.\*
- 31: /cgn2\_6/ptodata/2/paa/US105\_COMB.pcp.\*
- 32: /cgn2\_6/ptodata/2/paa/US106\_COMB.pcp.\*
- 33: /cgn2\_6/ptodata/2/paa/US107\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
------------	-------	--------------	----	-------------

1	3774	100.0	735	1	PCT-US03-35733-10	Sequence 10, Appl
2	3774	100.0	735	23	US-09-848-909A-10	Sequence 10, Appl
3	3771	99.9	735	1	PCT-US03-19786-4	Sequence 4, Appl
4	3768	99.8	735	1	PCT-US03-35733-2	Sequence 2, Appl
5	3768	99.8	735	23	US-09-848-909A-2	Sequence 2, Appl
6	3767	99.8	735	1	PCT-US03-35733-8	Sequence 8, Appl
7	3767	99.8	735	23	US-09-848-909A-8	Sequence 8, Appl
8	3766	99.8	735	1	PCT-US03-35733-13	Sequence 13, Appl
9	3766	99.8	735	23	US-09-848-909A-13	Sequence 13, Appl
10	3763	99.7	735	1	PCT-US03-35733-7	Sequence 7, Appl
11	3763	99.7	735	23	US-09-848-909A-7	Sequence 7, Appl
12	3762	99.7	735	1	PCT-US03-35733-4	Sequence 4, Appl
13	3762	99.7	735	23	US-09-848-909A-4	Sequence 4, Appl
14	3762	99.7	735	23	US-09-848-909A-6	Sequence 6, Appl
15	3761	99.7	735	1	PCT-US03-35733-5	Sequence 5, Appl
16	3761	99.7	735	1	PCT-US03-35733-11	Sequence 11, Appl
17	3761	99.7	735	1	PCT-US03-35733-19	Sequence 19, Appl
18	3761	99.7	735	1	PCT-US03-35733-20	Sequence 20, Appl
19	3761	99.7	735	1	PCT-US03-35733-21	Sequence 21, Appl
20	3761	99.7	735	22	US-09-791-537-43735	Sequence 43735, A
21	3761	99.7	735	23	US-09-848-909A-5	Sequence 5, Appl
22	3761	99.7	735	23	US-09-848-909A-11	Sequence 11, Appl
23	3761	99.7	735	23	US-09-848-909A-19	Sequence 19, Appl
24	3761	99.7	735	23	US-09-848-909A-20	Sequence 20, Appl
25	3761	99.7	735	23	US-09-848-909A-21	Sequence 21, Appl
26	3761	99.7	735	30	US-10-410-647-30	Sequence 30, Appl
27	3761	99.7	736	1	PCT-US01-14372A-1	Sequence 1, Appl
28	3761	99.7	736	1	PCT-US01-14372A-2	Sequence 2, Appl
29	3761	99.7	736	1	PCT-US01-14372A-3	Sequence 3, Appl
30	3761	99.7	736	1	PCT-US01-14372A-4	Sequence 4, Appl
31	3761	99.7	736	1	PCT-US01-14372A-5	Sequence 5, Appl
32	3761	99.7	736	1	PCT-US01-14372A-6	Sequence 6, Appl
33	3761	99.7	736	1	PCT-US01-14372A-7	Sequence 7, Appl
34	3761	99.7	736	1	PCT-US01-14372A-8	Sequence 8, Appl
35	3761	99.7	736	1	PCT-US01-14372A-9	Sequence 9, Appl
36	3761	99.7	736	1	PCT-US01-14372A-10	Sequence 10, Appl
37	3761	99.7	736	1	PCT-US01-14372A-11	Sequence 11, Appl
38	3761	99.7	736	1	PCT-US01-14372A-12	Sequence 12, Appl
39	3761	99.7	736	1	PCT-US01-14372A-13	Sequence 13, Appl
40	3761	99.7	736	1	PCT-US01-14372A-14	Sequence 14, Appl
41	3761	99.7	736	1	PCT-US01-14372A-15	Sequence 15, Appl
42	3761	99.7	736	1	PCT-US01-14372A-16	Sequence 16, Appl
43	3761	99.7	736	1	PCT-US01-14372A-17	Sequence 17, Appl
44	3761	99.7	736	1	PCT-US01-14372A-18	Sequence 18, Appl
45	3761	99.7	736	1	PCT-US01-14372A-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
PCT-US03-35733-10  
; Sequence 10, Application PC/TUS0335733  
; GENERAL INFORMATION:  
; APPLICANT: President and Fellows of Harvard College et al.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; of Bacterial Infection  
; FILE REFERENCE: 00742/072003  
; CURRENT APPLICATION NUMBER: PCT/US03/35733  
; CURRENT FILING DATE: 2003-10-10  
; PRIOR FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
PCT-US03-35733-10

Query Match 100.0%; Score 3774; DB 1; Length 735;  
Best Local Similarity 100.0%; Pred. No. 1.1e-299;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPISENQYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPISENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKOKSSNSRKRSTASGPTVPDRN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKOKSSNSRKRSTASGPTVPDRN 180  
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEYV 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEYV 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300  
QY 301 SEVHGNAEVAHGFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHGFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNNPENGVRVDTGSNWSEV 480  
Db 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNNPENGVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGFNEPGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGFNEPGNL 540  
QY 541 QYQKDIITFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKNAXMILIRDKRPHYDR 600  
Db 541 QYQKDIITFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKNAXMILIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 2

US-09-848-909A-10  
; Sequence 10, Application US/09848909A

; GENERAL INFORMATION:

; APPLICANT: Collier, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909A

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-05-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

## US-09-848-909A-10

Query Match 100.0%; Score 3774; DB 23; Length 735;  
Best Local Similarity 100.0%; Pred. No. 1.1e-299;  
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPISENQYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFAQPMVVTSTTGDLSIPSELENIPISENQYF 60  
QY 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSAADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120  
QY 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKOKSSNSRKRSTASGPTVPDRN 180  
Db 121 QRENTEKGLDFKLYWTDSONKKEVSSDNLQLPKOKSSNSRKRSTASGPTVPDRN 180  
QY 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEYV 240  
Db 181 DGPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPYDFEYV 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300  
Db 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKNTSRTHT 300  
QY 301 SEVHGNAEVAHGFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVAHGFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
Db 361 NANIRYVNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNNPENGVRVDTGSNWSEV 480  
Db 421 LNAQKDFSTPTIMYNNQFLEKTKQLRLDQVYGNIAIYNNPENGVRVDTGSNWSEV 480  
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGFNEPGNL 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSPLETTKPDMTLKEALKIAFGFNEPGNL 540  
QY 541 QYQKDIITFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKNAXMILIRDKRPHYDR 600  
Db 541 QYQKDIITFDFNFDOQTSONIKNQLAELNATNIYTVLDKIKNAXMILIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
Db 601 NNIAVGADESUVKEAHREVINSSTGLLNIDKIRKILSGYVIEIDTEGLKEVINDRY 660  
QY 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720  
Db 661 DMLNSSLRODQKTFIDFKKYNKPLIYISNPNYKVNVAVTKENTIINPSENGDSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 3

PCT-US03-19786-4

; Sequence 4, Application PC/TUS0319786

; GENERAL INFORMATION:

; APPLICANT: Human Genome Sciences, Inc.

; TITLE OF INVENTION: Modified Shine Dalgarno Sequences and Methods of Use Thereof

; FILE REFERENCE: PV595PCT

; CURRENT APPLICATION NUMBER: PCT/US03/19786

; CURRENT FILING DATE: 2003-06-25

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Artificial sequence		; PRIOR FILING DATE: 2002-11-08	
; FEATURE:		; NUMBER OF SEQ ID NOS: 38	
; OTHER INFORMATION: Mature PA sequence including an ETV signal sequence		; SOFTWARE: FastSeq for Windows Version 4.0	
PCT-US03-19786-4		; SEQ ID NO 2	
		; LENGTH: 735	
		; TYPE: PRT	
		; ORGANISM: Bacillus anthracis	
		PCT-US03-35733-2	
Query Match		99.8%; Score 3768; DB 1; Length 735;	
Best Local Similarity		99.9%; Pred. No. 3.3e-299;	
Matches 734; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	EVKQENLLNESSESSOGLGYYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF	60
Db	1	EVKQENLLNESSESSOGLGYYFSDLNFOAPMVVTSSTTGDLSIPSELENIPSENOYF	60
Qy	61	QSAIWSGFIKVKKSDYTFATSDAHVHTMWVDDQEVINKASNKIRLEKRLYQIKIY	120
Db	61	QSAIWSGFIKVKKSDYTFATSDAHVHTMWVDDQEVINKASNKIRLEKRLYQIKIY	120
Qy	121	QRENPTKGLDPKLYWTDSONKKEVIVSSDNLOLPKOKSSNRKSTASGTPVDPDRN	180
Db	121	QRENPTKGLDPKLYWTDSONKKEVIVSSDNLOLPKOKSSNRKSTASGTPVDPDRN	180
Qy	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYSDPEKVT	240
Db	181	DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWSSTASDPYSDPEKVT	240
Qy	241	GRIDKNVSPERHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
Db	241	GRIDKNVSPERHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT	300
Qy	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Db	301	SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL	360
Qy	361	NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPPSKNLAPIA	420
Db	361	NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPPSKNLAPIA	420
Qy	421	LNAQKDFSTPTTMNYNQFLEKTKQLRLDQVYGNIAIYFNENGRVVRVDTGSNWSEV	480
Db	421	LNAQKDFSTPTTMNYNQFLEKTKQLRLDQVYGNIAIYFNENGRVVRVDTGSNWSEV	480
Qy	481	LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAPGFNEPNGNL	540
Db	481	LPQIQETTARIIFNGKDLNLVERRIAAVNPSPDLETTKPDMTLKEALKIAPGFNEPNGNL	540
Qy	541	QYQGKDIETFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKQNTLIRDKRPHYDR	600
Db	541	QYQGKDIETFDNFQDQTSQNIKNQLAELNATNIYTVLDKIKLNKQNTLIRDKRPHYDR	600
Qy	601	NNIYVAGDESUVKEAHREVINSTEGLLLNIDKQIRKILSGYIIVEIEDTEGLKEVINDRY	660
Db	601	NNIYVAGDESUVKEAHREVINSTEGLLLNIDKQIRKILSGYIIVEIEDTEGLKEVINDRY	660
Qy	661	DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG	720
Db	661	DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIINPSENGDTSTNG	720
Qy	721	IKKILIFSKKGYEIG 735	
Db	721	IKKILIFSKKGYEIG 735	
RESULT 4			
PCT-US03-35733-2			
; Sequence 2, Application PC/TUS0335733			
; GENERAL INFORMATION:			
; APPLICANT: President and Fellows of Harvard College et al.			
; TITLE OF INVENTION: Compounds and Methods for the Treatment			
; FILE REFERENCE: 00742/072003			
; CURRENT APPLICATION NUMBER: PCT/US03/35733			
; CURRENT FILING DATE: 2003-10-10			
; PRIOR APPLICATION NUMBER: US 60/424,987			

```

; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; PRIOR FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-2

```

```

Query Match      99.8%; Score 3768; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 3.3e-299;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60

QY 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKCRSTAGTPVDRDN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKCRSTAGTPVDRDN 180

QY 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPSDFEKT 240
DB 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPSDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKTSTSRHT 300

QY 301 SEVHGNAEVHAGFFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHAGFFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAIYVNFENGVRVDTGSNWSEV 480
DB 421 LNAQDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAIYVNFENGVRVDTGSNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540

QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNKONILLIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNKONILLIRDKRPHYDR 600

QY 601 NNTAVGADESVMKEAHREVINSGTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
DB 601 NNTAVGADESVMKEAHREVINSGTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660

QY 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

```

RESULT 6  
PCT-US03-35733-8

```

; Sequence 8, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-8

```

```

Query Match      99.8%; Score 3767; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 3.9e-299;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60
DB 1 EVKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGDLSIPSSSELENIPSENYF 60

QY 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFIKVKSDYEYFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120

QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKCRSTAGTPVDRDN 180
DB 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPELKOKSSNRKCRSTAGTPVDRDN 180

QY 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPSDFEKT 240
DB 181 DGIPDSLEVEGYTVDVQNKRTFLSPWISNIHEKGLTKYKSSPEKWSASTASDPSDFEKT 240

QY 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKTSTSRHT 300
DB 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRISKTSTSRHT 300

QY 301 SEVHGNAEVHAGFFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHAGFFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

QY 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420
DB 361 NANIRVNTGTAPIYVNLPTTSLVLGKQNTLATIKADENQLSQILAPNNYPSKNLAPIA 420

QY 421 LNAQDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAIYVNFENGVRVDTGSNWSEV 480
DB 421 LNAQDFSSPTITMNYNQFLEKTKQLRDLTDQVYGNIAIYVNFENGVRVDTGSNWSEV 480

QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPNGNL 540

QY 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNKONILLIRDKRPHYDR 600
DB 541 QYQKDIITEFDNFDOQTSONIKQLAELNATNIYVLDKIKLNKONILLIRDKRPHYDR 600

QY 601 NNTAVGADESVMKEAHREVINSGTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660
DB 601 NNTAVGADESVMKEAHREVINSGTEGLLNIDKIRKILSGYIVIEIDTEGLKEVINDRY 660

QY 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720
DB 661 DMLNLSLRQDGKTFIDFKKYNKDLPLYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720

QY 721 IKKILIFSKKGYEIG 735
DB 721 IKKILIFSKKGYEIG 735

```

```
RESULT 7
US-09-848-909A-8
; Sequence 8, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Sellman, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; FILING DATE: 2000-05-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-8

Query Match          99.8%; Score 3767; DB 23; Length 735;
Best Local Similarity 99.9%; Pred. No. 3.9e-299;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKKSDDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFTKVKKSDDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRITSKNTSSTRHT 300
DB 241 GRIDKNVSPKARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRITSKNTSSTRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKQLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKQLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVRVDTGSNWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
DB 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660
QY 661 DMLNISSLRQDGKTFIDFKKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNISSLRQDGKTFIDFKKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
```

```
DB 661 DMLNISSLRQDGKTFIDFKKYNDKLPYISNPNYKVNVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIFSCKGYEIG 735
DB 721 IKKILIFSCKGYEIG 735

RESULT 8
PCT-US03-35733-13
; Sequence 13, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; TITLE OF INVENTION: and Prevention of Bacterial Infection
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; PRIOR FILING DATE: 2002-11-08
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-13
```

```
Query Match          99.8%; Score 3766; DB 1; Length 735;
Best Local Similarity 99.9%; Pred. No. 4.8e-299;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFTKVKKSDDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
DB 61 QSAIWSGFTKVKKSDDEYTFATSADNHVTMWVDDQEVINKASNSKIRLEKGRLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNSRKKRSTSGAGTVPDRDN 180
QY 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
DB 181 DGIPDSLEVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWSASDPYSDFEKT 240
QY 241 GRIDKNVSPKARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRITSKNTSSTRHT 300
DB 241 GRIDKNVSPKARHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRITSKNTSSTRHT 300
QY 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVAHSPFDIGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKQLAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKQLAPIA 420
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVRVDTGSNWSEV 480
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNATYFNENGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDPLETTKPDMTLKEALKIAGFNEPENG 540
QY 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
DB 541 QYQKDIETEDFDFDQSTQSNIKNQLAELNATNIYTVLDKIKLNAKNILIRDKRPHYDR 600
QY 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660
DB 601 NNIAVGADESVMKEAHEVINSSTEGLLNIDKIDKILSGYIVEIEDTEGLKEVINDRY 660
```

Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
 Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNKYKVVAVTKENTIIINPSENGDTSTNG 720  
 Db 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNKYKVVAVTKENTIIINPSENGDTSTNG 720  
 Qy 721 IKKILIFSKGVEIG 735  
 Db 721 IKKILIFSKGVEIG 735

RESULT 9  
 US-09-848-909A-13  
 ; Sequence 13, Application US/09848909A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Collier, R. John  
 ; APPLICANT: Sellman, Brett R.  
 ; TITLE OF INVENTION: Compounds and Methods for the Treatment  
 ; TITLE OF INVENTION: and Prevention of Bacterial Infection  
 ; FILE REFERENCE: 00742/060002  
 ; CURRENT APPLICATION NUMBER: US/09/848,909A  
 ; PRIOR FILING DATE: 2001-05-04  
 ; PRIOR APPLICATION NUMBER: US 60/201,800  
 ; PRIOR FILING DATE: 2000-05-04  
 ; NUMBER OF SEQ ID NOS: 35  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13  
 ; LENGTH: 735  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus anthracis  
 US-09-848-909A-13

Query Match 99.8%; Score 3766; DB 23; Length 735;  
 Best Local Similarity 99.9%; Pred. No. 4.8e-299;  
 Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPISENOYF 60  
 Db 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPISENOYF 60

Qy 61 QSAIWSGFIKVKKSDDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120  
 Db 61 QSAIWSGFIKVKKSDDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKRSSTAGTVPDRDN 180  
 Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKRSSTAGTVPDRDN 180

Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYDFEKT 240  
 Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
 Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRYVNTGTAPINYVLTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420  
 Db 361 NANIRYVNTGTAPINYVLTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGRVVDTSNWSSEV 480  
 Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGRVVDTSNWSSEV 480

Qy 481 LPOIQTETARIIFNGKOLNVERIAAVNPSDELETTKPDWMTKEALKIAGFNEPNGNL 540  
 Db 481 LPOIQTETARIIFNGKOLNVERIAAVNPSDELETTKPDWMTKEALKIAGFNEPNGNL 540

Qy 541 QYQKDIETEPFNDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600

Db 541 QYQKDIETEPFNDQOTSQNIKNQLAELNATNIYTVLDKIKLNAKNNILIRDKRFHYDR 600  
 Qy 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
 Db 601 NNIAVGADES VVKEAHEVINSSTEGLLNIDKDIRKILSGYIVEIEDTEGLKEVINDRY 660  
 Qy 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNKYKVVAVTKENTIIINPSENGDTSTNG 720  
 Db 661 DMLNISSLRQDGKTFIDFKYNDKPLYSINPNKYKVVAVTKENTIIINPSENGDTSTNG 720  
 Qy 721 IKKILIFSKGVEIG 735  
 Db 721 IKKILIFSKGVEIG 735

RESULT 10  
 PCT-US03-35733-7  
 ; Sequence 7, Application PC/TUS0335733  
 ; GENERAL INFORMATION:  
 ; APPLICANT: President and Fellows of Harvard College et al.  
 ; TITLE OF INVENTION: Compounds and Methods for the Treatment  
 ; TITLE OF INVENTION: and Prevention of Bacterial Infection  
 ; FILE REFERENCE: 00742/072003  
 ; CURRENT APPLICATION NUMBER: PCT/US03/35733  
 ; CURRENT FILING DATE: 2003-10-10  
 ; PRIOR APPLICATION NUMBER: US 60/424,987  
 ; PRIOR FILING DATE: 2002-11-08  
 ; NUMBER OF SEQ ID NOS: 38  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 735  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus anthracis  
 PCT-US03-35733-7

Query Match 99.7%; Score 3763; DB 1; Length 735;  
 Best Local Similarity 99.7%; Pred. No. 8.4e-299;  
 Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPISENOYF 60  
 Db 1 EVKQENRLNSESSESSQGLLYFSDLNFPQPMVVTSSITGDLSPSSSELENIPISENOYF 60

Qy 61 QSAIWSGFIKVKKSDDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120  
 Db 61 QSAIWSGFIKVKKSDDEYTFATSDNHNVTWVDDQEVINKASNNKIRLEKGLYQIKIY 120

Qy 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKRSSTAGTVPDRDN 180  
 Db 121 QRENPTKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKRSSTAGTVPDRDN 180

Qy 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYDFEKT 240  
 Db 181 DGI PDSLEVEGYTVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSSTASDPYDFEKT 240

Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
 Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSRTHT 300

Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360  
 Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360

Qy 361 NANIRYVNTGTAPINYVLTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420  
 Db 361 NANIRYVNTGTAPINYVLTSLVLGKNQTLATIKADENQSLQILAPNNYPSKNLAPIA 420

Qy 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGRVVDTSNWSSEV 480  
 Db 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDTPQVYGNATYNFENGRVVDTSNWSSEV 480

Qy 481 LPOIQTETARIIFNGKOLNVERIAAVNPSDELETTKPDWMTKEALKIAGFNEPNGNL 540

Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDLETTKPDMLKALKIAFGNEPENG 540  
Qy 541 QYQCKDITFDFNFDQOTSONIKQNLAEALNATNIYVLDKIKLNAKWNILIRDRKRFHYDR 600  
Db 541 QYQCKDITFDFNFDQOTSONIKQNLAEALNATNIYVLDKIKLNAKWNILIRDRKRFHYDR 600  
Qy 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Qy 661 DMLNSSLRODQGTFTDFKKNYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQGTFTDFKKNYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 11

US-09-848-909A-7

; Sequence 7, Application US/09848909A

; GENERAL INFORMATION:

; APPLICANT: Sellman, R. John

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/060002

; CURRENT APPLICATION NUMBER: US/09/848,909A

; CURRENT FILING DATE: 2001-05-04

; PRIOR APPLICATION NUMBER: US 60/201,800

; PRIOR FILING DATE: 2000-05-04

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-09-848-909A-7

Query Match 99.7%; Score 3763; DB 23; Length 735;  
Best Local Similarity 99.7%; Pred. No. 8.4e-299;  
Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSNRKRSTSGAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSNRKRSTSGAGTPVDRDN 180  
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240  
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360  
Qy 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLPIA 420  
Db 361 NANIRYNTGTAPYINVLPTTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLPIA 420

Qy 421 LNAQKDPSTPTIMYNOFLEKTKQRLDQVYGNIAFYENGSRVVRTGNSMSEV 480  
Db 421 LNAQKDPSTPTIMYNOFLEKTKQRLDQVYGNIAFYENGSRVVRTGNSMSEV 480  
Qy 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDLETTKPDMLKALKIAFGNEPENG 540  
Db 481 LPOIQTETARIIFNGKDLNVERRIAANVPSDLETTKPDMLKALKIAFGNEPENG 540  
Qy 541 QYQCKDITFDFNFDQOTSONIKQNLAEALNATNIYVLDKIKLNAKWNILIRDRKRFHYDR 600  
Db 541 QYQCKDITFDFNFDQOTSONIKQNLAEALNATNIYVLDKIKLNAKWNILIRDRKRFHYDR 600  
Qy 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Db 601 NNTAVGADESUVKEAAREVINSSTEGLLNIDKDIRKILSGYIVIEIDTEGLKEVINDRY 660  
Qy 661 DMLNSSLRODQGTFTDFKKNYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
Db 661 DMLNSSLRODQGTFTDFKKNYNDKPLIYISNPNYKVNVAVTKENTIIINPSENGDTSTNG 720  
Qy 721 IKKILIFSKKGYEIG 735  
Db 721 IKKILIFSKKGYEIG 735

## RESULT 12

PCT-US03-35733-4

; Sequence 4, Application PC/TUS0335733

; GENERAL INFORMATION:

; APPLICANT: President and Fellows of Harvard College et al.

; TITLE OF INVENTION: Compounds and Methods for the Treatment

; TITLE OF INVENTION: and Prevention of Bacterial Infection

; FILE REFERENCE: 00742/072003

; CURRENT APPLICATION NUMBER: PCT/US03/35733

; CURRENT FILING DATE: 2003-10-10

; PRIOR APPLICATION NUMBER: US 60/424,987

; PRIOR FILING DATE: 2002-11-08

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 735

; TYPE: PRT

; ORGANISM: Bacillus anthracis

PCT-US03-35733-4

Query Match 99.7%; Score 3762; DB 1; Length 735;  
Best Local Similarity 99.7%; Pred. No. 1e-298;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
Db 1 EVKQENRLNSESQGLGYFSDLNFPQAPMVVTSSTTGDLSIPSSSELENIPSENOYF 60  
Qy 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120  
Db 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNKIRLEKGLYQIKIY 120  
Qy 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSNRKRSTSGAGTPVDRDN 180  
Db 121 QRENPTKGLDFKLYWTDSONKKEVSSDNQLQPELKQKSNRKRSTSGAGTPVDRDN 180  
Qy 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240  
Db 181 DGIIPDSLEVEGYTVDVKNKRTFLSPWISNHEKGLTKYKSPKWSSTASDPYDFEKT 240  
Qy 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Db 241 GRIDKNVSPPEARHPLVAAYPIVHVDMENTILSKNEDQSTQNTDSETRTISKNTSRTHT 300  
Qy 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360  
Db 301 SEVHGNAEVHASFDDIGGSVAGFSNSNSTVAIDHSLAGERTWAETMGLNTADTARL 360

```

QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKULAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKULAPIA 420
QY 421 LNAQKOPSSPTITMYNOFLEKTKQLRLDQVYGNATYFNGRVRVDTGSNWSEV 480
DB 421 LNAQKOPSSPTITMYNOFLEKTKQLRLDQVYGNATYFNGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLTKEALKIAFGNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLTKEALKIAFGNEPENG 540
QY 541 QYQKDIETEDFNFDOOTSQNIKNQAEINATNIYVLDKIKLNKAMNILLRDKRFHYDR 600
DB 541 QYQKDIETEDFNFDOOTSQNIKNQAEINATNIYVLDKIKLNKAMNILLRDKRFHYDR 600
QY 601 NNTAVGADESUVKEAUREVINSSTEGLLNIDKIRKILSGYIVIEBTEGLKEVINDRY 660
DB 601 NNTAVGADESUVKEAUREVINSSTEGLLNIDKIRKILSGYIVIEBTEGLKEVINDRY 660
QY 661 DMLNSSLRODGKTFIDFKYNDKPLIYISNPYKVVAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNSSLRODGKTFIDFKYNDKPLIYISNPYKVVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIPSKGYEIG 735
DB 721 IKKILIPSKGYEIG 735

RESULT 13
PCT-US03-35733-6
; Sequence 6, Application PC/TUS0335733
; GENERAL INFORMATION:
; APPLICANT: President and Fellows of Harvard College et al.
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/072003
; CURRENT APPLICATION NUMBER: PCT/US03/35733
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US 60/424,987
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; TYPE: PRT
; ORGANISM: Bacillus anthracis
PCT-US03-35733-6

```

```

Query Match 99.7%; Score 3762; DB 1; Length 735;
Best Local Similarity 99.7%; Pred. No. 1e-298;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKGSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKGSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
QY 181 DGIPODSEVEGYTVDVVKRKTFTLSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240
DB 181 DGIPODSEVEGYTVDVVKRKTFTLSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240
QY 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSEPTTSKNTSRTHT 300
DB 241 GRIDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSEPTTSKNTSRTHT 300

```

```

QY 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
DB 301 SEVHGNAEVHASFDFDGGSVSAGFSNSNSTVAIDHSLSLAGERTWAETMGLNTADTARL 360
QY 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKULAPIA 420
DB 361 NANIRYVNTGTAPIYVNLPTTSLVGLKQNTLATIKADENQLSQILAPNNYPSKULAPIA 420
QY 421 LNAQKOPSSPTITMYNOFLEKTKQLRLDQVYGNATYFNGRVRVDTGSNWSEV 480
DB 421 LNAQKOPSSPTITMYNOFLEKTKQLRLDQVYGNATYFNGRVRVDTGSNWSEV 480
QY 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLTKEALKIAFGNEPENG 540
DB 481 LPOIQTETARIIFNGKDLNVERRIAAVNPSPLETTKPDMLTKEALKIAFGNEPENG 540
QY 541 QYQKDIETEDFNFDOOTSQNIKNQAEINATNIYVLDKIKLNKAMNILLRDKRFHYDR 600
DB 541 QYQKDIETEDFNFDOOTSQNIKNQAEINATNIYVLDKIKLNKAMNILLRDKRFHYDR 600
QY 601 NNTAVGADESUVKEAUREVINSSTEGLLNIDKIRKILSGYIVIEBTEGLKEVINDRY 660
DB 601 NNTAVGADESUVKEAUREVINSSTEGLLNIDKIRKILSGYIVIEBTEGLKEVINDRY 660
QY 661 DMLNSSLRODGKTFIDFKYNDKPLIYISNPYKVVAVTKENTIINPSENGDTSTNG 720
DB 661 DMLNSSLRODGKTFIDFKYNDKPLIYISNPYKVVAVTKENTIINPSENGDTSTNG 720
QY 721 IKKILIPSKGYEIG 735
DB 721 IKKILIPSKGYEIG 735

RESULT 14
US-09-848-909A-4
; Sequence 4, Application US/09848909A
; GENERAL INFORMATION:
; APPLICANT: Collier, R. John
; TITLE OF INVENTION: Compounds and Methods for the Treatment
; FILE REFERENCE: 00742/060002
; CURRENT APPLICATION NUMBER: US/09/848,909A
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/201,800
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-848-909A-4

```

```

Query Match 99.7%; Score 3762; DB 23; Length 735;
Best Local Similarity 99.7%; Pred. No. 1e-298;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
DB 1 EVKQENRLNSESSESSQGLLYFSDLNFOAPMVVTSSTTGLSIPSSSELENIPSENQYF 60
QY 61 QSAIWSGFIKVKGSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
DB 61 QSAIWSGFIKVKGSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIY 120
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQPELKOKSSNRKRSSTAGTVPDRDN 180
QY 181 DGIPODSEVEGYTVDVVKRKTFTLSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240
DB 181 DGIPODSEVEGYTVDVVKRKTFTLSPWISNIHEKKGLTKYKSPKWTASDPSPEKVT 240

```

QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL 360  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV 480  
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV 480  
QY 481 LPQIQETTARIIFNGKOLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
DB 481 LPQIQETTARIIFNGKOLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
QY 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR 600  
DB 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660  
DB 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKYNDKLPYISPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
DB 661 DMLNSSLRQDGKTFIDFKKYNDKLPYISPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

RESULT 15  
US-09-848-909A-6  
; Sequence 6, Application US/09848909A  
; GENERAL INFORMATION:  
; APPLICANT: Collier, R. John  
; APPLICANT: Sellman, Brett R.  
; TITLE OF INVENTION: Compounds and Methods for the Treatment  
; TITLE OF INVENTION: and Prevention of Bacterial Infection  
; FILE REFERENCE: 00742/060002  
; CURRENT APPLICATION NUMBER: US/09/848,909A  
; CURRENT FILING DATE: 2001-05-04  
; PRIOR APPLICATION NUMBER: US 60/201,800  
; PRIOR FILING DATE: 2000-05-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 735  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-09-848-909A-6

Query Match 99.7%; Score 3762; DB 23; Length 735;  
Best Local Similarity 99.7%; Pred. No. 1e-298;  
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EVKQENRLNSESSESSQGLGYFFDLNFQAPMWVTSSTGDLSPSSELENIPSENQYF 60  
DB 1 EVKQENRLNSESSESSQGLGYFFDLNFQAPMWVTSSTGDLSPSSELENIPSENQYF 60  
QY 61 QSAIWSGFIKVKSDSYTTFATADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120  
DB 61 QSAIWSGFIKVKSDSYTTFATADNHVTWVDDQEVINKASNSKIRLEKGRLYQIKQY 120  
QY 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQIPKQKSSNRKRKSTAGTVPDRDN 180

DB 121 QRENPTKGLDFKLYWTDSONKKEVSSDNLQIPKQKSSNRKRKSTAGTVPDRDN 180  
QY 181 DCIPDSLEVEGYVDVKNKRTFLSPWISNIIHEKKGLTKYKSSPEKMWSTASDPYSDFEKT 240  
DB 181 DCIPDSLEVEGYVDVKNKRTFLSPWISNIIHEKKGLTKYKSSPEKMWSTASDPYSDFEKT 240  
QY 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
DB 241 GRIDKNVSPARHPLVAAYPIVHVDMENIILSKNEDQSTQNTDSETRTISKNTSTSRHT 300  
QY 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL 360  
DB 301 SEVHGNAEVAHSPFDIGGSVAGFSNSNSTVAIDHSLSLAGERWAEWTGLNTADTARL 360  
QY 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
DB 361 NANIRYVNTGTAPIYNNVLPPTSLVLGKNQTLATIKADENQLSQILAPNNYPSKNLAPIA 420  
QY 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV 480  
DB 421 LNAQKDFSSPTITMNYNQFLEKTKQLRLDQVYGNIAATYVNFENGVRVDTGNSWSEV 480  
QY 481 LPQIQETTARIIFNGKOLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
DB 481 LPQIQETTARIIFNGKOLNLVERRIAAVNPSDPLETTKPDMLKEALKIAFGNEPNGNL 540  
QY 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR 600  
DB 541 QYQKDIITEFDNFDOQTSQNIKNQLAELNATNIYTVLDKIKLAKNMLIRDKRPHYDR 600  
QY 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660  
DB 601 NNIAVGADESUVKEAHREVINSSTEGLLNIDKDIRKILSGYIIVEIEDTEGLKEVINDRY 660  
QY 661 DMLNSSLRQDGKTFIDFKKYNDKLPYISPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
DB 661 DMLNSSLRQDGKTFIDFKKYNDKLPYISPNYKVNVAVTYKNTIINPSENGDTSTNG 720  
QY 721 IKKILIFSKKGYEIG 735  
DB 721 IKKILIFSKKGYEIG 735

Search completed: May 3, 2004, 19:57:40  
Job time : 178.546 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:30:47 ; Search time 13.0855 Seconds  
(without alignments)  
5403.004 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENLLNESSSQGL.....TSTNGIKLIFSKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3761	99.7	764	2 I39934	protective antigen
2	876	23.2	875	2 I40862	iota toxin compone
3	235.5	6.2	192	2 I39933	cryptic protein -
4	235.5	6.2	204	2 G59104	hypothetical prote
5	206	5.5	4688	2 F82885	hypothetical prote
6	188.5	5.0	2401	2 T28676	rhostry protein -
7	185.5	4.9	1639	2 S05603	major merozoite su
8	184.5	4.9	1125	2 E95998	membrane nuclease,
9	184.5	4.9	2529	2 B64635	toxin-like outer m
10	182	4.8	6713	2 B89921	hypothetical prote
11	181	4.8	2869	2 T28677	rhostry protein -
12	181	4.8	4152	2 T31102	filamentous hemagg
13	180	4.8	1193	2 S68218	botulinum neurotox
14	179.5	4.8	1072	2 A68627	hypothetical prote
15	179	4.7	1635	2 A10452	hemolysin [importe
16	178	4.7	1227	2 C97033	uncharacterized pr
17	177	4.7	752	2 G90599	hypothetical prote
18	177	4.7	1302	1 JC6009	surface-located me
19	177	4.7	4319	2 T31105	hypothetical prote
20	174.5	4.6	1365	2 T30822	lmp1 protein - Myc
21	174	4.6	1658	2 S55101	repeat organellar
22	173	4.6	1939	2 T18372	hypothetical prote
23	173	4.6	5005	2 F82884	toxin-like outer m
24	172.5	4.5	2399	2 H71879	hypothetical prote
25	171.5	4.5	821	2 S67087	alpha-toxin - C1os
26	171	4.5	2178	2 S55805	major merozoite su
27	170	4.5	1631	1 SAZQK1	probable nuclear p
28	170	4.5	1837	2 T41023	SMC1 protein homol
29	169.5	4.5	1230	2 S56850	

ALIGNMENTS

RESULT 1

I39934

protective antigen precursor - Bacillus anthracis plasmid

C;Species: Bacillus anthracis

C;Date: 19-Jul-1996 #sequence revision 19-Jul-1996 #text\_change 01-Dec-2000

C;Accession: I39934; S89160; F59104

R;Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppa, S.H.; Schmidt, J.J.  
Gene 69, 287-300, 1988

A;Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr.

A;Reference number: I39933; MUID:89172073; PMID:3148491

A;Accession: I39934

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-764 <RES>

A;Cross-references: GB:M25589; NID:9143280; PIDN:AA22637.1; PID:9143282

R;Friedman, T.C.; Gordon, V.M.; Leppa, S.H.; Klmpel, K.R.; Birch, N.P.; Loh, Y.P.  
Arch. Biochem. Biophys. 316, 5-13, 1995

A;Title: In vitro processing of anthrax toxin protective antigen by recombinant PC1 (SP

A;Reference number: S69160; MUID:95142670; PMID:7840657

A;Accession: S69160

A;Molecule type: protein

A;Residues: 197-202 <FRI>

R;Okinkaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehle

J. Bacteriol. 181, 6509-6515, 1999

A;Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbor

A;Reference number: A59091; MUID:99445483; PMID:10515943

A;Accession: F59104

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-313,'Q',315-764 <OKI>

A;Cross-references: GB:AF065404; NID:G4894216; PIDN:AA22414.1; PID:G4894326

A;Experimental source: strain Sterne

A;Note: similar to anthrax toxin moiety, protective antigen, pagA formerly pag, plasmid

C;Genetics:

A;Gene: pXO1-110

A;Genome: plasmid

C;Function:

A;Description: three component exotoxin; protective antigen binds to receptors on the s

Y active components edema factor or lethal factor; the complex is internalized by recep

C;Keywords: exotoxin

F;1-29/Domain: signal sequence #status predicted <SIG>

F;30-196/Domain: propeptide #status predicted <PRO>

F;197-202/Product: protective antigen #status experimental <MAT>

Query Match 99.7%; Score 3761; DB 2; Length 764;

Best Local Similarity 99.7%; Pred. No. 2.5e-182;

Matches 733; Conservativity 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EVKQENLLNESSSQGLGGYFSDLNFOAPMVVTSITGDLISPSSELENIPSENQYF 60

DB 30 EVKQENLLNESSSQGLGGYFSDLNFOAPMVVTSITGDLISPSSELENIPSENQYF 89

QY 61 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 120  
DB 90 QSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 149  
QY 121 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKKSSTAGTPVPRDN 180  
DB 150 QRENTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKKSSTAGTPVPRDN 209  
QY 181 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 240  
DB 210 DGIPOSLEVEGYTVVKNKRTFLSPWISNIHEKGLTKYKSPKSWSTASDPYDFEKT 269  
QY 241 GRIDKNVSPPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSTRTISKNTSTSRHT 300  
DB 270 GRIDKNVSPPEARHPLVAAPYVHVDMENIILSKNEDQSTQNTDSTRTISKNTSTSRHT 329  
QY 301 SEVHGNAEVSAGFNSNSTVAIDHSISLAGERTWAETMGLNTADTARL 360  
DB 330 SEVHGNAEVSAGFNSNSTVAIDHSISLAGERTWAETMGLNTADTARL 389  
QY 361 NANIRVYNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKULAPIA 420  
DB 390 NANIRVYNTGTAPIYNNVLTSLVLGKQNTLATIKADENQLSQILAPNNYPSKULAPIA 449  
QY 421 LNAQKDFSSPTIMYNNQFLEKTKQLRLDQVYGNATYNNFENGVRVDTGSNWSEV 480  
DB 450 LNAQKDFSSPTIMYNNQFLEKTKQLRLDQVYGNATYNNFENGVRVDTGSNWSEV 509  
QY 481 LPQIQETTARIIFNGKDLNVERRAAVNPSDPLETTKPDMLKEALKIATGFNEPNCNL 540  
DB 510 LPQIQETTARIIFNGKDLNVERRAAVNPSDPLETTKPDMLKEALKIATGFNEPNCNL 569  
QY 541 QYQKDIITPDPNFDQSTQNTKQNLAEANATYVLDKIKLNAKQILIRDKKPHYDR 600  
DB 570 QYQKDIITPDPNFDQSTQNTKQNLAEANATYVLDKIKLNAKQILIRDKKPHYDR 629  
QY 601 NNIAVGADESVEKAEHREVNSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 660  
DB 630 NNIAVGADESVEKAEHREVNSTEGLLNIDKIRKILSGYVIEDETEGLKEVINDRY 689  
QY 661 DMLNISLRQDKTDFDKKNDKPLIYTSNPNYKNTVAVTKENTIINPSENGTSTNG 720  
DB 690 DMLNISLRQDKTDFDKKNDKPLIYTSNPNYKNTVAVTKENTIINPSENGTSTNG 749  
QY 721 IKKILIFSKGYEIG 735  
DB 750 IKKILIFSKGYEIG 764

RESULT 2  
I40862  
iota toxin component Ib - Clostridium perfringens  
C:Species: Clostridium perfringens  
C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 15-Oct-1999  
C:Accession: I40862; S42774  
R:Perelle, S.; Gibert, M.; Boquet, P.; Popoff, M.R.  
Infect. Immun. 61, 5147-5156, 1993  
A:Title: Characterization of Clostridium perfringens iota-toxin genes and expression in  
A:Reference number: I40861; MUID:94041637; PMID:8225592  
A:Accession: I40862  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-875 <RES>  
A:Cross-references: EMBL:X73562; NID:g929031; PIDN:CAA51960.1; PID:g414655  
Query Match 23.2%; Score 876; DB 2; Length 875;  
Best Local Similarity 31.2%; Pred. No. 1.2e-36;  
Matches 253; Conservative 130; Mismatches 278; Indels 150; Gaps 31;  
QY 1 EVKQENRLANESSSQGLGYFSDUNFQAPMVVTSSTGYDUSIPSSLENIIPSP-NQY 59  
DB 35 DTNQKEBITNTLSSNGLMGYFADBFKDLLEMAPKNGDLKFKKXVDKLLTDEDS 94

QY 60 FOSAIWSGFIKVKKSDEYTFATSDADNHVTWVDDQEVINKASNSKIRLEKGLYQIKIQY 119  
DB 95 IKSIRWTGRIIPSEDEGYILSTDR-NDVLQINAKGDIKAK---TLKVNKKGOAYNIRIE 150  
QY 120 YQREN-----PTEKGLDFKLYWTDSONKKEVISSDNLQLPKQKSSNRKKSST 170  
DB 151 IODKNLGSIDNLVSP-----KLW-ELUNGKNTVPEENLFFRDYKSIDEND----- 195  
QY 171 AGTVP-----DRNDGIPDSLEVEGYTVVKNKRTFLSPWISNIH 211  
DB 196 --PFINNFFDVRFPSAAWEDEDLTDNDNIPDAVEKNGYTI---KDSIAVKWNSFA 249  
QY 212 EKGLTKYKSSPKWSTASDPYDFEKTGRIDKNVSPPEARHPLVAAPYVHVDMENIIL 271  
DB 250 E-QYKRYVSSYLESTAGDPYTDYKASGSDKAKLEARDPLVAAPYVGVGMENLII 308  
QY 272 SKVEDOSTQNTDSTRTISKNTSTSTHTSEVHGNAEVSAGFNSNST 331  
DB 309 STNEHASS---DQKTVSRATYNSKTANTV-----GVSIAGYQNGFTGN 351  
QY 332 VAIDHS-----LSLAGERTWAETMGLNTADTARLNAIRYNTGTAPIYNNVLTSL 383  
DB 352 ITTSYSHTTDNSTAVQDSNGESMNTGLSINKGESAVINANVRYNTGTAPYKVTPTNL 411  
QY 384 VLKGNQTLATIKADENQLSQILAPNNYPSKULAPIALNAQKDFSTPTIMYNNQFLE 443  
DB 412 VL-DGETLATIKADENQILGNNLSPNETYPKKGLSPALNTMQFNARLIPINYDQLKLD 470  
QY 444 KTKQLRLDQVYGNATYNNFENGVRVDTGSNWSEVLPQIOETTARIIFNGKDLNVER 503  
DB 471 SGKQIKLETTQVSGNYGTXN-SQGI-ITEGNSWNYISQIDSVSASIIID-TGSQTFER 527  
QY 504 RIAAVNPSDPLETTKPDMLKEALKIATGFNEPNCNLQY-QGKDIIE--PDFNFDQSTQ 560  
DB 528 RVAASQGNPEDKT-PEITIGEAIKAFSATK-NGELLYFNGIPIDESCVELIPDDNTSE 585  
QY 561 NIKNQLAELNATYVLDKIKLNAKQILIRDKRF-HYDR-NNIAVGADESVEKAEH 617  
DB 586 IIEQLKYLDDKKIYXV---KLERGNLILKVPSTFTNFDEYNFP--ASWSNIDTKNQ 639  
QY 618 EVINSSTEGT-----LLNIDKIRKILSGY-----IYEIETDEGLKEVIND 658  
DB 640 DGLQSVANKLSETKIIPMSKLPYKRYVFGSYGKDPSTSNITVNIKSKQKTDYLPV 699  
QY 659 RYDMLNIS-----SLRQDKTDFDKKNDKPLIYTSNPNYKNTVAVTKENTIINP 696  
DB 700 EKDYTKESYEFETTKGKSSDIEITLSSGVIFLDNLISITELNSTPEILKEPEIKVPSDOE 759  
QY 697 -----NVYAVTKENTIINPSENGTSTNGI 721  
DB 760 ILDAHKKYADIKLDT-----NTGNTYIDGI 785

RESULT 3  
I39933  
cryptic protein - Bacillus anthracis  
C:Species: Bacillus anthracis  
C>Date: 19-Jul-1996 #sequence\_revision 19-Jul-1996 #text\_change 15-Oct-1999  
C:Accession: I39933  
R:Welkos, S.L.; Lowe, J.R.; Eden-McCutchan, F.; Vodkin, M.; Leppla, S.H.; Schmidt, J.J.  
Gene 69, 287-300, 1988  
A:Title: Sequence and analysis of the DNA encoding protective antigen of Bacillus anthr  
A:Reference number: I39933; MUID:89172073; PMID:3148491  
A:Accession: I39933  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-192 <RES>  
A:Cross-references: GB:M22589; NID:g143280; PIDN:AAA23636.1; PID:g143281

Query Match 6.2%; Score 235.5; DB 2; Length 192;  
Best Local Similarity 34.6%; Pred. No. 2.6e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNTLIRKPHDRNNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIVEI 646  
 Db 1 MNTLVRDP-YHYDNNNGNIIVGVDDSVLKAYKQILNWSGVSGLNLDVQALSGYMLQI 59  
 QY 647 EDTE-----GLKEVINDRYDMLNITSSLRQDGKTFIDPKYNDKLPYISNPN 693  
 Db 60 KPSNHLTNSPVTITLAGKDSGVGELYRVL-----DAGFLDNKFDENWRSIV-DPG 112  
 QY 694 YKNVYAVTKEN-TIINPSENGDSTNGIKKILIFSKGYEI 734  
 Db 113 DDVYVAVTKEDFPNAVTRDENGNI-NKLKNTLVLSGKIKEI 153

RESULT 4  
 G59104  
 hypothetical protein pXOI-111 - Bacillus anthracis virulence plasmid pXOI  
 C/Species: Bacillus anthracis  
 C/Date: 12-Nov-1999 #sequence\_revision 12-Nov-1999 #text\_change 11-May-2000  
 C/Accession: G59104  
 R/Okinaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler  
 J. Bacteriol. 181, 6509-6515, 1999  
 A/Title: Sequence and organization of pXOI, the large Bacillus anthracis plasmid harbored  
 A/Reference number: A59091, MUID:99445483; PMID:10515943  
 A/Accession: G59104  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-204 <OKI>  
 A/Cross-references: GB:AF065404; NID:g4894216; PIDN:AA032415.1; PID:g4894327  
 A/Experimental source: strain Sterne  
 A/Note: similar to hypothetical protein in the protective antigen domain; ypa, plasmid p  
 C/Genetics:  
 A/Gene: pXOI-111  
 A/Genome: plasmid

Query Match 6.2%; Score 235.5; DB 2; Length 204;  
 Best Local Similarity 34.6%; Pred. No. 2.8e-05;  
 Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNTLIRKPHDRNNIAGADESVVKEAHRVINSSTEGLLNIDKIRKILSGYIVEI 646  
 Db 1 MNTLVRDP-YHYDNNNGNIIVGVDDSVLKAYKQILNWSGVSGLNLDVQALSGYMLQI 59  
 QY 647 EDTE-----GLKEVINDRYDMLNITSSLRQDGKTFIDPKYNDKLPYISNPN 693  
 Db 60 KPSNHLTNSPVTITLAGKDSGVGELYRVL-----DAGFLDNKFDENWRSIV-DPG 112  
 QY 694 YKNVYAVTKEN-TIINPSENGDSTNGIKKILIFSKGYEI 734  
 Db 113 DDVYVAVTKEDFPNAVTRDENGNI-NKLKNTLVLSGKIKEI 153

RESULT 5  
 F82885  
 hypothetical protein U0482 [imported] - Ureaplasma urealyticum  
 C/Species: Ureaplasma urealyticum  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C/Accession: F82885  
 R/Glass, J.I.; Leikowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 submitted to GenBank, February 2000  
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
 A/Reference number: A82870  
 A/Accession: F82885  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-468 <GLA>  
 A/Cross-references: GB:AE002145; GB:AF22894; NID:g5699476; PIDN:AAF30894.1; GSPDB:GN001  
 A/Experimental source: serovar 3; biovar 1  
 C/Genetics:  
 A/Gene: U0482  
 A/Genetic code: SGC3

Query Match 5.5%; Score 206; DB 2; Length 468;  
 Best Local Similarity 21.5%; Pred. No. 0.09;

Matches 174; Conservative 132; Mismatches 324; Indels 186; Gaps 40;  
 QY 19 LLGYVFSOLFQAPMVVTSSTTGLDLSIPSSSELENIPESENQVFOASIGWGIKVKKSDYIT 78  
 Db 3699 LVDVYLD-NHONIDEIRKIFKOHNV-SKEIENPGVTWISKGNWKSSTDTITANEFK 3756  
 QY 79 PATSADNHVWVDOEVINVASNSKIRLEKGLYQIKIQORENTEKGLDFKLTWD 138  
 Db 3757 IETQ-----DDNDVLNIDATVRFKDEHNNIKQIKVRIKEN-----ND 3795  
 QY 139 SQNKKEVSSDNLQPELKQSSN-----SRKRTSAGTPVDRD----- 179  
 Db 3796 WLKQOI---DNLN-PETKYKLENIELSKPLKTHNLVSINDKENISLITGCPVLKV 3851  
 QY 180 -----NDGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSPKWSASDPSD 235  
 Db 3852 IQTQNDTINDTQQTINVTLSGNSK-YNGROIKVYKDNNNVYESS---LITLQKQND 3907  
 QY 236 FEKVTRGRIDKQVSPPEARPLVAAPIVHVDMENIILSKN-EDOSTQTDSET-RTISKNT 293  
 Db 3908 YQLLSNLNSN-----REYRFEKIEINHSINTNNFEDLEKLVGVSNTFITQTKT 3957  
 QY 294 STSRTHTS-EVHGNAEVASP-----FDIGSVSAGPS-----NSNS---STVAD 335  
 Db 3958 TVQWNDSSATVGTGVNFNFKIKSEDKILENNOQVAVFAPKETIRDTTWLQYTRPLK 4017  
 QY 336 HSLSLAGERTWAETMGLNT---ADTARLNANIRYVNTGTAPIYVNLPTTSLVLGKQTLA 392  
 Db 4018 DVTSPFKSGTWAHDLNSVNFKEETTYKLVKIQFVKNPTRAKONNINNNVILDTNSI 4077  
 QY 393 -----TIKADENQLSILAPNNYPPSKLAPIALN-AOKDFSTPTWNY--NQFLE 443  
 Db 4078 NENYEFTTKVGDHKLINITSNNVNTNSQTINFTLSGVKKSWGKKTKLSYKNDSESI 4137  
 QY 444 KTKQLRLTDQVYGNIAFYNGR-----VRVDTGSNMSEVLPTQIETARIIFNGKD 497  
 Db 4138 HNEVLIESNKTQYVILLNLLNLRKRTVTLIDVKLIDNNVSDPPKEGNTLNSFITRTSA 4197  
 QY 498 LNLVEERAAVNPSPLETT-----KPDWTLKEA-----LKIAGFNEPENGLO 541  
 Db 4198 INVLNIEISNRASNTLSKSTIIKINLNDPQVLRDKQDQATIVYGNKKQAMGFTVSGNIK 4257  
 QY 542 YQKQDITDFDFDQDQTSQNTK-NQLAELNATNTYVLDKIKLNKAKNIIIRDKRHYD- 599  
 Db 4258 YLTATLVLDNFN-DKVNIVNISFNKKPSIAEN-----IGDKSNII-----YNNDS 4304  
 QY 600 -----RNIIAAGA---DESVVKEAHRVINSSTEGLLNIDKIRKILSGYIVEIDTE 650  
 Db 4305 IPKLEINNDIIVGPIKKEIVVKNQK--NNIDVDLGLQINPKIAHLR-FIAKFKSTN 4361  
 QY 651 GLKEVIND--RYDMLNITSSLRQDGKTFIDF-----KKY----- 681  
 Db 4362 -----NDIETNVINGSSLVNNDGKTSIRFTLNNLKANKLYSLVDVYVLVNNNSNTIVE 4415  
 QY 682 NDKLPYISNPNYKNVYAVTKENTIINPSENGDTS 717  
 Db 4416 SNKLP-KLNNINYQIK---INKSHTII--SKNGEWS 4445

RESULT 6  
 T28676  
 rhoxy protein - plasmodium yoelii (fragment)  
 C/Species: Plasmodium yoelii  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
 C/Accession: T28676; A45521  
 R/Sinha, K.A.; Keen, J.K.; Ogun, S.A.; Holder, A.A.  
 Mol. Biochem. Parasitol. 76, 329-332, 1996  
 A/Title: Comparison of two members of a multigene family coding for high-molecular mass  
 A/Reference number: Z20507; MUID:97077455; PMID:8920022  
 A/Accession: T28676  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-2401 <SIN>

A/Cross-references: EMBL:U36927; NID:G1041784; PID:G1041785; PIDN:AAB41263.1  
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.  
Mol. Biochem. Parasitol. 42: 241-246, 1990  
A/Title: Identification of the gene for a Plasmodium yoelii rhothry protein. Multiple cd  
A/Reference number: A45521; MUID:91101660; PMID:2270106  
A/Accession: A45521  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 2260-2401 <KEE>  
A/Cross-references: GB:M34281

Query Match 5.0%; Score 188.5; DB 2; Length 2401;  
Best Local Similarity 21.7%; Pred. No. 0.26; Indels 221; Gaps 43;  
Matches 182; Conservative 119; Mismatches 316;

QY 1 EVKQENRLNSESSESSQGLGYFSDLN-----FQAPMVVTSSTGDLGI 45  
DB 402 EVAKENVQNVYKSNILKEIKHYNDQINIDNIKEKAKQNDQPKHEMKTIPTTPNEMKYQK 461  
QY 46 PSSELENIPSNQVQSAL--WSGFIKY---KKSDEYTFATSDNHNVTMWVDDQEV---I 97  
DB 462 PSIEIKMKDE---FLSKVKNYNDPDKYKSEHNKFTLTNKTIVSDSEIKKYE 518  
QY 98 NKASNSKIRLEKGRLYQIKYQORENTEKGLDFKLYWTDSONKKEVIGSDNLQLPELK 157  
DB 519 NKFNDSKLNETKK--SIESEYQINILKKVDD--YIKVCLNTNELITNCHNKQITLK 573  
QY 158 QKSNRK--KRSTAGTVPDRDNDGPD-----SLEVEGYTVDVKNRTFLSP 205  
DB 574 DKLNQNIKTETNSIDKIYTDKFNILTDKKLETKFTGLSLNNHESNKELLTYFYD 633  
QY 206 WSNTHKEKGLTKYKSSPEKSTASDPYSDPEKVTGRIDKNVSPPEARPLVAAPVHVVD 265  
DB 634 LKANIQKKNMLYQFNEK---EKAVEDIKKQVNDKLVSN-----LEIT 677  
QY 266 MENILSKNEDQSTONQSTETISKNSTSTSTHTSEVHGNAEVHASFDDIGGSVAGFS 325  
DB 678 IVTSIYNED-----TENE---IGKSIELNLTNKLVE-----KYKANVT 713  
QY 326 NSNSSTVAI--DHSLSLAGRTWAETMGLTADTASLNANIRVNTGTAPIVNVLPITSLV 384  
DB 714 NLNEIKKADQDFQFGK-----EKNIKYPDEN--KIKNDIDTLNQK 754  
QY 385 LGKN--OTLATIKAD--ENQLSQI-----LAPN---NYYP---SKNLAPIALNAQK-- 425  
DB 755 IDKSIETLTTEIKGNSHDEIKGQIDKLVKVPNTMFNEDPKETEKIENIVEKIDKKK 814  
QY 426 -----DFSTPTVMYNOFLEKTKQLRLDQVYGNIAVYFENGVRVDTGNSWSE 479  
DB 815 NIYKEIDKULNEISKIENDKTSLEKJNINLSYKSLGNLFLOQIDEEKKAEHTIKAME 874  
QY 480 V---LPOIQTETARIIFNGKDLN-----VERRIAAVNPSDP-----LETT-----K 518  
DB 875 AYIDDLNLIKKSCEI---EKENVINMDIKMDIHKEMKALNISHDDYKIYHTTSKNHEEK 931  
QY 519 PMTWLKEALKIAGFNEP---NGNLOYQKQDTEFPDFNFQDQTSQNIKNQLAELNATNIY 575  
DB 932 ISDIRKNSLKIIDPFSESYINDIKKELEKNVLE-----SONNNTDINGYLSKIE--NIY 984  
QY 576 TVLDKIKLNAMILIRDKRF--HYDRNNIAGADESVVKEAHREVINSGTEGLLNIDK 633  
DB 985 NIL---KLNKIKIIDKVEYTDIEKN-----KKINAEIENS-----1020  
QY 634 DIRKILSGYIVETEDTGLKE-----VINDRY-----DMLNSSLROQDKTIDPKYND 683  
DB 1021 --EKI-----ITOLKENSLLKCEOSKIKSTIDDNVSECKNTNL-----KTVIVNEKN-- 1069  
QY 684 KLPILYISN--PNYKVN-----YAVTKENTIINSENGDTSN--GIKKILIFSCK 730  
DB 1070 -INTYFKNAZEYQNVSLNFNEMADTKSQILNLIKKNNGTNTNIDYNIKELKHKKK 1126

RESULT 7

S05603

major merozoite surface antigen precursor - malaria parasite (Plasmodium falciparum) (s  
N/Alternate names: gp195 surface antigen  
C/Species: Plasmodium falciparum  
C/Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 09-Jun-2000  
C/Accession: S05603; S04850

R/Mylar, P.J.

submitted to the EMBL Data Library, April 1989

A/Reference number: S05603

A/Accession: S05603

A/Molecule type: mRNA

A/Residues: 1-1639 &lt;MYL&gt;

A/Cross-references: EMBL:X15063; NID:G9896; PIDN:CAA33163.1; PID:G9897

R/Mylar, P.J.

Nucleic Acids Res. 17, 5401, 1989

A/Title: Nucleotide and deduced amino acid sequence of the gp195 (MSA-1) gene from Plas  
A/Reference number: S04850; MUID:89345116; PMID:2668887

A/Accession: S04850

A/Molecule type: mRNA

A/Residues: 1504-1639 &lt;MYL2&gt;

A/Cross-references: EMBL:X15063

C/Species: major merozoite surface antigen

C/Keywords: glycoprotein; merozoite; surface antigen

F;1-19/Domain: signal sequence #status predicted &lt;SIG&gt;

F;20-1639/Product: major merozoite surface antigen #status predicted &lt;MAT&gt;

Query Match 4.9%; Score 185.5; DB 2; Length 1639;  
Best Local Similarity 20.5%; Pred. No. 0.21;  
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;QY 5 ENRL-LNESESSQGLGYFSDLNFOAPMVVTSSTGDLSPSSELENIPSENQYFOSA 63  
DB 927 ENTILSGKNYQELIGQKSE-NF-----YKILKDSSTFTYNE 965  
QY 64 IWGFIKVKKSBYTFATSDNHNVTMWVDDQVINKASNS-----NKIRLEKGRLY- 114  
DB 966 SFTNFVKSADD-----INSLNDESKKLEEDINKKLTQLSPDLYNKYKLERLFD 1020  
QY 115 -----QIKIQORENPTKGLDFKLYWTDSONKKEVIGSDNLQLPELKQKSSNRKKKS 168  
DB 1021 KKTVGKQKQIKKLLKEQLESKL--NSLANPKHVL--QNFSVFFNKKAEAEIAETEN 1076  
QY 169 TSAGTVPDRDNDGIPDSLEVEGYTVDVKNKRTFTSPWISNIHEKKGLTKY---KSSP-- 223  
DB 1077 T-----LENTKILLGHY-----KGLVKNYNGESSPLK 1103  
QY 224 ---EKWSTASDPYSDPE-----KVTGRIDKNVSPPEAR-----HPLVAAPVHVVD 265  
DB 1104 TLSEESIQTEDNYASLENFKVLSKLEGKLDNLEKKLSYLSGLHLLIA-----E 1156  
QY 266 MENILSKNEDQSTONQSTETISKNSTSTSTHTSEVHGNAEVHASFDDIGGSVAGFS 325  
DB 1157 LKEVIRKKN---YTGNSPSNNP-----DVNNALSYKFFLEPGTDVATVVS 1200  
QY 326 NSNSSTVAIDHSLSLAGRTWAETMGLTADTASLNANIRVNTGTAPIVNVLPIT- 381  
DB 1201 ESGSDTLEQSQPKKPASTHVGAES--NTITTSQ-NVDDDEVDDVIVIPGESEEDYDDL 1256  
QY 382 -SLVLGKNQTLATKADENQLSQILAPNYYPSKNLAPIA---LNAAKDFSSPTITWYVN 437  
DB 1257 GQVVTGEAVTPSVI---DNILSKI---ENEYEVLYLKLAVGTVRSKKQLENNVMTFNVN 1310  
QY 438 -----QFLELEKTKQLRLDQV-YGNIAI-----YFENGVRVDTGNSWSEVL 481  
DB 1311 VKDILNSRFNKNENFKNV-LESDDLIPYKDLTSSNYVVKDPYKFLNKKRDKFLSSVNYIK 1369  
QY 482 PQIQTETARIIFNGKDLNVERRIAANVPSDPL-----ETTKPM-TLKEALKIAGF 533  
DB 1370 DSID-----TDINFA-----NDVLGYKYLSEKYSKSDLSIKKYNKQGE 1410  
QY 534 NEPN-----GNLOYQKQDTE-----PDFNFQDQTSQNIKNQLAELNATNIYT 576  
DB 1411 NEKLPFLNNIETLYKTVNDKIDLFIHLEAKVLYNYTEK---SNVEVKIKELN--YLKT 1465

QY 577 VLDKIKLNKAKN--ILIRDKFHYDRNN-----IAGV-ADESUVKEAHRVINSSTEGLL 628  
Db 1466 IODKLADFKKNNFVGIADLTSDYNNHLLTKFLSTGMVFENLAKTVLSLNDLQNG-M 1524  
QY 629 LNIIDKD--IRKILSGYVIEIDTSLKLEVDNDRYDMLNLSISLQDQGTTFIDFKKYNKLP 686  
Db 1525 LNIHQHCVK-----OCQNSGCTFRHDE-----RECKCLINYQEGDKC- 1566  
QY 687 LYISNPYKNVYAVTKEINTIINPSENG-----DTSNIGIKI 724  
Db 1567 --VENPNTCN-----EN-----NGGCDADAKTEEDSGSNG-KKI 1599

RESULT 8  
E90598  
membrane nuclease, lipoprotein [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C:Species: Mycoplasma pulmonis  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
C:Accession: E90598  
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galissou, F.; Moerzer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
A:Reference number: A99512; MUID:21267165; PMID:1133084  
A:Accession: E90598  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1125 <KUR>  
A:Cross-references: GB:AU445566; PID:gl4090108; PIDN:CAC13866.1; GSPDB:GN00153  
A:Experimental source: strain UAB CTIP  
C:Genetics:  
A:Gene: MYPV 6930  
A:Genetic code: GCG3

Query Match 4.9%; Score 184.5; DB 2; Length 1125;  
Best Local Similarity 19.1%; Pred. No. 0.13;  
Matches 165; Conservative 162; Mismatches 330; Indels 207; Gaps 40;  
QY 3 KOENRLNSESOGLLGYFDFLNFORPMVVTSTTGDLSPSELENT:PSNQYFQS 62  
Db 120 ENDNSVNNKENSCK-----NDEN-----LTLTKVRLGHNV-LNQSGNIPTKMLAISK 166  
QY 63 A1WS-----GFIKV-----KASDRTTATSADNHTVWVDQD-----VI 97  
Db 169 VILHNKLDIVGLTEITNENGVKTIINELNKSDDKMYIVLSLKKGGTTGSGQBEHVGI 228  
QY 98 NKASNSNKIRLE-----KGLYQIKIQORENPTKGL----- 130  
Db 229 YK---ENKLTLSFDDKESKGFENKWL---DDPFKGGKIDFVRPPFGVKSTKGN 282  
QY 131 --DFKLYW--TDSQNKKE---VSSDNLQLPKQKSNRSKRSTAGTVPDRDNGI 183  
Db 283 KNDFTVFMHSDAPGVKEERGETSAGYSGQGHKEVAEALRTKEYNEYPDSI-----DGV 337  
QY 184 PDSLEVEGYT-----VVKNKRTFLSP-W--ISNIHEK---K 214  
Db 338 NNELFMGDTNKLGNKAPPLLOSQYKSLIKVKNATSLAORWEYAHYDKIYK 397  
QY 215 GLTKYKSSP--EKWSTASDPYSDFEKVTGRIDK-NVSPPEARHP-----LVAAYPI 261  
Db 398 GDLKVENSGFYDLWKVFDNLLNKEBFKVEYRSKKAQKDKYKEXSVYLHAI SDHTI 457  
QY 262 VHVDMENILSKNEDOSTQNTDSEITISKNTSTSEVHGNAEVAHSPFDIGGSVS 321  
Db 458 VYTDL--ILQKDD-----QNKSENK--DENSDSKQNDKQDNLSSKNDTPQSKSSPOI 509  
QY 322 AGFSNSNSTVAIDHSLSLAGERTWAETMGLTADTARLNANIRYVNTGTAPIYVLPPT 381  
Db 510 DASQNSNTTNEKDKQLSDODE---SKNNAIKSQNDQKSDNLSSKNDTPQSKSSPOI 566  
QY 382 SLVIGKQNTLTKADENQLSQILAPNYPYKSLAPIALNAQKDPSSPTITWYNQFLE 441  
Db 567 NPNLENNQSHSNGENDSDSKQNTSNRQTKN-----DLRSQKQKNTL----- 611

QY 442 LEKTKQLDLTDQVYGNATYNFENGVR---VDTGSNKNSEVLPOIQETTARIIFNGKDL 498  
Db 612 ---TKXPSNSNSNETKNETQNNENSSTKKDEIDTSA-----KTQDSTNSLNKNEBKT 661  
QY 499 NUVERRI-----AAVNPSPLETTKPDMLTKEALKI--AFGFNEPNGNLQYQKQDITE 549  
Db 662 NQVEKTNTESSNNSNTKQENSSSTKKEEISKSSESNNVNSSTNKQENIDNKKEEISK 721  
QY 550 FDFNFDQTSQNIKQ-LAELNATNIYTVLDKILNKNKN-----LLJEDKFFHYDRNI 603  
Db 722 SESNVNSNSTNTQNETPETNESQNNVIGKPNNSQNLQNAIDVSAKKVIGYWNINE 781  
QY 604 AVGADESUVKEAHRVINSSTEGLL---LNIIDKIRKILSGYVIEIDTGLKEVINDR 659  
Db 782 SVGKSASAKAFKAVKVIDHKNLDLVIGGLVHEETLTKI-----VEEMNKLKSDSDK 834  
QY 660 YDMLNLSLRQDQKTF-IDFKYNDKLPYISNPNYKNNVAVTKENT---IINSENGD 715  
Db 835 W--VQVISEKKQGEFPVNLARYIG--VIYKEN---KFNIESFKQNTNKGHLNENQPMN 887  
QY 716 TSTNGIKKI-----LIFSCKG 731  
Db 888 SSFNTSEKSVYVRPPFGIKFSTKG 911

RESULT 9  
B64635  
toxin-like outer membrane protein HP0922 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C>Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
C:Accession: B64635  
R:Tomb, J.F.; White, O.; Keriavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKen  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpik, P.D.; Smith, H.O.; Fraser, C  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467; PMID:9252185  
A:Accession: B64635  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2529 <TOM>  
A:Cross-references: GB:AE000602; GB:AE000511; NID:g2314060; PIDN:AAD07969.1; PID:g23140

Query Match 4.9%; Score 184.5; DB 2; Length 2529;  
Best Local Similarity 21.0%; Pred. No. 0.44;  
Matches 158; Conservative 94; Mismatches 261; Indels 241; Gaps 36;  
QY 123 ENPTEKGLDFKLYWTDSONK---KEVISSDNLQLPKQKSNRSKRSTAGTVPDR 178  
Db 729 QNPPAESVWSGVWL--QNKYSNKGIYID---PNLSQSGSGNTLSTYTANLF--- 779  
QY 179 DNDGIPDSLEVEGYTVVDKMKRTFLSP---WI-----SNHEKKGLTKYK 220  
Db 780 ---GRSFVNIQNTLIIGNNTSVNSGLIWIGHGFGYITGTFSAANIY-----LTNNF 832  
QY 221 SSEPKWT-----ASD-----PYSDFEKTGTGRIDKNVSPPEARHPVAAAYPIV 262  
Db 833 KTGEVNSGCGGANITFKASDNITMDGLNDAEIVTKMIQTGAS---QHSATFPALN 889  
QY 263 HVDMENTILSKNEDQSTQNTDSEITISKNTSTSRHTSEVHGNAEVAHSPFDIGGSVSA 322  
Db 890 NISVTNSFS-----DMTWGKFPSPKAKNISFS-----NASF-----S 921  
QY 323 GFSNSNSTVA---IDHSLSLAGERTWAETMGLTADTARLNANIRYVNTGTAPIYVNL- 378  
Db 922 GFTNPGSSVISANATNLSLFINR-----LNGGAVYLNQANSLIFN-NQAVFNVLY 973  
QY 379 -----PTSLVIGKQNTLTKADENQLSQILAPNYPYKSLAPIAL--NAQKDFS 428  
Db 974 SRGTSENFATTLQGLNTNFTLSS-----QSLNENFGDTTLQNNANITLGNKSOAFAK 1025



Db 485 QTKLENKFTFSLNHNEANNELI-KYFSDLKANLGINENMLYNQFTEKEKTFNDIKEX 543  
Qy 57 NOYFQSAIMSGFIKVKKSDRYTATADNHVTWVDD-----QEVINKASNSKIRLEK 110  
Db 544 NIHINEISKIEIKIHAS-IYNISEETEREIGINIESLTKVFVKENVTNLNIK-EK 601  
Qy 111 GRLYQIKIQORENPTEKGLDFLYWTDQNK-----KEVSSDNLQ----- 152  
Db 602 LKHDF-----SDFGKEGNIKYTDKIKINDIMAVSQIDOHINGLDIDQKSES 652  
Qy 153 -LPELKOKSNSRKRKSTASGTPVDRNDGIPDSLEVEGYTVDVKNKRTFLSPWISNIH 211  
Db 653 YVSEMEQINKLEKVSNT-----EISND-----NVEG-----IKKQII--VTKID 692  
Qy 212 EKKG-----LKYKSSPEKMTASDPYSDPEKVTG-----RIDKNVSPARPL 255  
Db 693 KKNIVBEINKLSBKIEKMTSLKVKVDNLSYGQNLGNLFQIDEE-KKAENTI 751  
Qy 256 VAAPIVHVDMENIILSKEDQSTQNTDSETRTISKNTSTSTHSEVHNAEVH---A 311  
Db 752 KSEAYID-DLDNKKKSQIEFEMDIKMDINKEMEALKISHDDDKCKDKSKHKNENIS 810  
Qy 312 SFDIGGSVAGPSNSSTVAIDHSLSLAGERWAEWTGLTADTARLNARIVYWTGT 371  
Db 811 DIDKSKIQDFSR-ESDINDIKNLQ-----KNVSSQHNDSINOCLNEV 857  
Qy 372 APIYVNLPTTSLVLGKNOTLATIKADENQSLAPNNYPS-----KNLAPIALNAQK 425  
Db 858 ANIYNILKL-----NKIKIIDKVKVETSEIEKNKKNINDELNNSEK 899  
Qy 426 -----DFSSPTITWYNOFLB-----LEKTKQLRLDTD-----QVYCN 458  
Db 900 VIKIEBGLSKCRKINSTLDDKIDECIKNINVLKKNILNEETNTNHPKNAEYK 959  
Qy 459 IATYFNGVRVDTGNSWSEVLPOIQTETARIIFNGKDLNLVERRIAAVPSDPLETTK 518  
Db 960 IVLNSFNN-----IEMADNKSQYLEIKKNGT---NDHDYNIKELK-SHKDKSNGYK-TE 1010  
Qy 519 PDWTLKALIAFGFNEPNGLVOYQKQD-----ITEFDNFDOOTSONIKQALAE 568  
Db 1011 ADQNKKAIAQ-----NKELEFQYKEEVULLNKYAVEULKNPD-KTKNDSQIITKE 1061  
Qy 569 LNATNIVTVLDKILNAKMN-----ILIRDRFRHYDRNNIAGDAESVVEKAHREVIN- 621  
Db 1062 IKAHNTCTLESQKSEKKNKNEIKHIEDEVANNDKSKAKITSIKVSEVPKTKIKI 1121  
Qy 622 ---SSTEGLLNIDKIRKILSGYIYEIEDT-----EGLEKVINDRYDMLNLSLR 669  
Db 1122 NEIRTKSDCLKETNLEKQISLSDTQETKLTENGKQKLTESL-----LESLEK 1172  
Qy 670 QDGKTFIDFKYNDKPLYSN-----PNYKVVYAVTKENTIINPSENGDTSTNG 720  
Db 1173 KQKKNIEDQKDELDEVNSKIKNTENTVNOHKKNYEIGI--VEKINEIAKTKNQIESTKE 1230  
Qy 721 IKK 723  
Db 1231 LIK 1233

## RESULT 12

T31102  
filamentous hemagglutinin 1 - Haemophilus ducreyi  
C/Species: Haemophilus ducreyi  
C/Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999  
C/Accession: T31102  
R/Ward, C.K.; Lumbley, S.R.; Latimer, J.L.; Cope, L.D.; Hansen, E.J.  
J. Bacteriol. 180, 6013-6022, 1998  
A/Title: Haemophilus ducreyi secretes a filamentous hemagglutinin-like protein.  
A/Reference number: Z20984; MUID:99030326; PMID:9811662  
A/Accession: T31102  
A/Status: preliminary; translated from GB/EMBL/DDBJ  
A/Molecule type: DNA  
A/Residues: 1-4152 <WAR>

A/Cross-references: EMBL:AF057695; MID:g3929017; PID:g3929018; PIDN:AAC79757.1  
C/Genetics:  
A/Gene: lspa1

Query Match 4.8%; Score 181; DB 2; Length 4152;  
Best Local Similarity 22.2%; Pred. No. 1.4; 321; Indels 218; Gaps 45;  
Matches 18; Conservative 118; Mismatches

Qy 10 NESSESSQGLGYFSDNLNQ-APMVVTSSTG-----LSIPSELENISENQYF 60  
Db 162 SRTESTSQ-LVGKLANIQKEAKLIILNQVTDGHSNIQGALEVAGKADLIIVN----- 216  
Qy 61 QSAIMSGFIKVKKSDRYTATADNHVTWVDDQVINKASN-----SNKIRLEKGRLY 114  
Db 217 PNCITLNGVKTINTDRFVVSST-----DIIPRENGLLSVNGKVTIDKGGVA 264  
Qy 115 QIKIQY---ORENPTEKGLDFKLYWTDSONKKEVSSD-----NLQPELKOKKS 161  
Db 265 TNGLSHFVEVARNIDQKG---KITVAKTENQKSNVNPANITFAAGSLNLTNLTREATPSS 321  
Qy 162 NSRKKESTSAGTVPDRNDGIPDSLEVEGYTVD---VKNKRTPLSPWISNIHEKKGLT 217  
Db 322 GTSR---TSDTFAI-SAGSAGSMYGSNIKFFVTDKGAGVKHGIIFSENDINIKMDGNA 377  
Qy 218 KYKSSPEKSTASDPYSDPEKVTGIDKNVGEARHPLVAAPIVHVDMENIILSKEDQ 277  
Db 378 SLKELYAK-----KQIDILAKDIELTEKGOLOANNKIILNSTGKINLRNASEV 425  
Qy 278 STQNTDSETRTIS-KNTSTSRHTSEVHNAEVAHSPFDIGGSVAGPSNSSTVAIDH 336  
Db 426 SADNVNKSLENALENASMSANSLDVITKLEVRNS-----SKVSGATNIAKSNITLDG 480  
Qy 337 SLSLAGERTW-----AETWGLNTADTARLN-----ANIRVYN---T 369  
Db 481 SSVANKITLNTVNTNATLNQSKLSAKOMELNVTHNITLNTSKLSAQKANKIKENTLN 540  
Qy 370 GTAPIVNLPTTSLVLGK-----NOTLA-----TIKADENO-LSOILAPNN-- 409  
Db 541 GEA-----SLVAEKLDINAIDKIYNTNGTIAGLTANITTKALENDRNADLIILAHQNLN 591  
Qy 410 -----YY-----PSKNLAPIALNAQKDFSS-----TPITMNVNQPLETEKTKQL 448  
Db 592 FTVNGSHYNNKGDIVSKCKAIVTFSSNDSFTSNGSKLVDAQNNLTNVNVTNITQSGEII 651  
Qy 449 RLDTDOVYGNI---ATYAPEN-GRVVDTSNGWSEVLPOIQTETARIIFNGKDLNLVER- 503  
Db 652 -----LHGVNVLNAKGNFTNSGNLTTKMLNINIESFI---NAGNLTGKNLEVHSNT 702  
Qy 504 -----RIAAVPSDPLETTKEDMTL-KEALKIAPG-----FNEPKNLOYOQKDIETEF 550  
Db 703 TVKNDGKLYSIEENLNISSKTDFTNNGTLLGLEALKIASGNGFTNASNGSLA-SNKSLDIY 761  
Qy 551 DENF-DQQTSONIKKQALAEALNATNIYTVLDK--IKLNAKMNILIRDRFRHYDRNNIAGVA 607  
Db 762 GNNFTNGTIESVKS-----LMTNNYTFINNATIKSYGLVNLITSQG-NFTNDSNGTWSH 816  
Qy 608 D-ESVVKEAHREVINSSST-----EGULL-----NIDKIRKILSGYIYEIEDTEGLKEVIN 658  
Db 817 DLINITSQAN--IINKLLAGGQGLNLTAKGNITNDSNSTAIAVLHNSNDIN-----LNA 869  
Qy 659 RYDMLNLSL-RQDGKTFIDFKYNDKPLYSNPNYKVVYAVTKENTIINPSENGDTS 717  
Db 870 NKKYVNIIGIYSQAGNISVEAKLHNDVKL-----SGNITTTK-----SGNATVK 915  
Qy 718 TNGI 721  
Db 916 TNSI 919

## RESULT 13

S68218  
botulinum neurotoxin type A, nontoxic/nonhemagglutinating - Clostridium botulinum (stra  
C/Species: Clostridium botulinum

A: Variety: strain NIH  
C: Date: 15-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 02-Jun-2003  
C: Accession: S68218; S74301  
R: Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.  
FEBS Lett. 376. 41-44. 1995  
A: Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components  
A: Reference number: S67988; MUID: 96096783; PMID: 8521962  
A: Accession: S68218  
A: Molecule type: DNA  
A: Residues: 1-1193 <FUJ>  
A: Cross-references: EMBL: D67030; NID: g2160224; PIDN: BAA11050.1; PID: d1011710; PID: g11324  
A: Experimental source: strain NIH  
A: Accession: S74301  
A: Molecule type: protein  
A: Residues: 1-13,145-155 <FUI>  
A: Experimental source: strain NIH  
C: Genetics:  
A: Gene: ant  
C: Superfamily: tetanus toxin  
C: Keywords: neurotoxin  
F: 1-1193/Product: botulinum neurotoxin type A, nontoxic/nonhemagglutinating #status expected

Query Match 4.8%; Score 180; DB 2; Length 1193;  
Best Local Similarity 19.6%; Pred. No. 0.24;  
Matches 181; Conservative 131; Mismatches 319; Indels 294; Gaps 46;

QY 42 DLSIP---SSELENIPSENQYFOSAIWGFIRKVK--KSDBYTPATS-----ADNHVT 88  
DB 241 DLVTPYLRSELENIYSQNLNIVDLVSGGIDPKFINTDPYFTDNYFSAKVPEDHRN 300  
QY 89 MWVDQEVINKASNSKIRLEK-----GRLYQIKIQY-----QRENPTEKGLDF 132  
DB 301 IYETEIEGNAIGNDIKLRLKQKFRININDIWEINLNYFSKFSIMWPRFNALKHFYR 360  
QY 133 KLYW-----TDSQNKKEVISS----- 148  
DB 361 QYKYKIDYPENYSINGFVNGQINLAQLSLSDRNQDILNKPEIINLLNGNVNLSRNIYV 420  
QY 149 -----DNQLP-----ELKQKSSNRKRSTSGPTVPDRDNDGIDPSLEVEGY 192  
DB 421 DGLKSTVDDFVSNYKIPNRYAYEHFNNSDLSLDNNGIY-----DNIEIDVNPY 474  
QY 193 TVD-----VKNKRTFLS--PWISNIHEKGLTKYKSSPEKWSASTADPYSDPEKVTG 241  
DB 475 KENCDFKFPVKITSTREINTNIPWPNY-----LQAQNTNNEKFSLS-----SDFVEVVS 525  
QY 242 RIDKNSVPPEARHPLVAAPYVHVDMENII--LSKNEDQSTQNTDSE-----TRTISKNYST 295  
DB 526 SKDKS-----LVYFSLNVMYFLDSIKDINSPIIDTKYIYLWLEIFRNYGF 571  
QY 296 SRTHSTSEVHNAEVH-----ASFDDIGGSVSAFGNSNS--STVAIDHSLSLAGERT 345  
DB 572 DITATQETINTCGINKVVTWFGKALNLTNTSD3FVEEFQNLGAISLKNKENLSMPLIES 631  
QY 346 WA---ETWGLNTADTARLANIRVYNTG--TAPIYVNLPT-----TSVLGKQNT 390  
DB 632 YEIPNDMLUGLPINDLNEKLFNIYSKNTAYPKKYVYNFLDQWWTQYYSQYFDDICWAKRSV 691  
QY 391 LA---TIKADENQLSOLAPNYPYKXNLAPIAL---NAQKPFSS--TPITMN-----435  
DB 692 LAQETLIKRIIQKLSVLIGNSISDNLALMNLTTNTLRDISNESQIAMANVDSFLNN 751  
QY 436 -----NQFLELEKTKQLRLDQVYG--NIATYFNGRVRVDTGSMSEVLBPQ 483  
DB 752 AAICVFESNIYPKFISF-----NEQCINNINIKTEF-----IQKCTNINEDEKUL 798  
QY 484 IOETTARIIFNGKDLNVERIAAVNPSDFLETKPDMTLKEAL-----KIAFGFNEPNGN 539  
DB 799 INQN-----VFNSLDPEFLN-----IQNKSLFSSETALLIKEETWPELVLYAFKEPGN 849  
QY 540 -----LOYOGKDI--TEPDFNFD-----QOTSQNIKNQALNATNIYTV---577  
DB 850 VIGDASGKNTSIEY--SKDIGVYINGSDALVNGSNQSI5FSNDFFENGITNSFSIYFWL 908

QY 578 --LDKIKLNAKM-----NILIRD---KPFY--- 598  
DB 909 RNLGKDTIKSLGSKEDNCGWEIFQDTGLVFNMDSGNKEKNYLSGVNNSWHYITI 968  
QY 599 --DR--NNIYAGADESVV--KEAHRVINSSTEGILLNIDKDIRKILSGYIVIEIEDTEGLK 653  
DB 969 SVDRLEKQULLIFIDDLNLANVESIKELINLIYSNIIISLSENNPSYIEGTLINLPKTS-Q 1027  
QY 654 EVINDRYDMLNIISSFLRQDKTPIDFKYNKNDKPLIYSNPYKNV---YVAVTKENTI--- 707  
DB 1028 EVLSNYFEVLNNSYIYSDNEERLE---YNYKTYLY---NVVFSKPICEVKQNNIYLT 1080  
QY 708 INPSENGDSTNGIKKILIFSKGY 732  
DB 1081 INNTNLLNQASKFKLLSNPNKQY 1105

RESULT 14  
A: 86827  
hypothetical protein ynfG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C: Species: Lactococcus lactis subsp. lactis  
C: Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C: Accession: A86827  
R: Bolotin, A.; Wincker, P.; Mager, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrl  
Genome Res. 11. 731-753. 2001  
A: Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s.  
A: Reference number: A86625; MUID: 21235186; PMID: 11337471  
A: Accession: A86827  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-1072 <STO>  
A: Cross-references: GB: A8005176; PID: g12724625; PIDN: AAK05715.1; GSPDB: GN00146  
A: Experimental source: strain IL1403  
C: Genetics:  
A: Gene: ynfG

Query Match 4.8%; Score 179.5; DB 2; Length 1072;  
Best Local Similarity 19.4%; Pred. No. 0.22;  
Matches 160; Conservative 142; Mismatches 342; Indels 179; Gaps 30;

QY 2 VKQENRLNSESSESSQGLGYFSDLNFOAPMVVTSSTTGLSLSPSELEN-----52  
DB 300 VEQVDVASSESTQDANSALYPISEASSVTDNLANSISLSDSISSQTSQSGASSTA 359  
QY 53 --IPSENQYFQSA---IWSGFIKVKSDYETFATSDNHNVTMWYDDQEVINKASNSKI 106  
DB 360 EISYDSENSLSLSSNQINSNSKDSNQSLGSSMSSESEHSNSNSINETNNSSEI 419  
QY 107 R-----LEKRLYQIKIQYQRENPTKGLDFKL-----YWTDQNKKEVISSDNQLPELK 157  
DB 420 TNILPSPNPTESNSVSDQTSSEASTNSNSISLSLSPNISSTSDRSATNSDFSNVAEVA 479  
QY 158 QKSSNRKRKSTSGPTVPDRNDGI-----PDSLEVEGYTVDKNKRTPLSWNIHEK 213  
DB 480 NNSLASVNNSSSVLSSTSTADNLGINSQSNLTKD---SSEISSTGAFLS---SNQTS 534  
QY 214 KGLTKYS---SPEKWS-----TASDPYSDPEKVTGRIDKNVSPPEARHPLVAAPYI 261  
DB 535 EASTNSNSISLSLSPNISSTVLESTSSNFSNFAEVANNGLASVNNSSSVLSSTSTA 594  
QY 262 VHYDM-----ENI-----ILSKNEQDQSTQNTDSETRTISKNTSTRTSTSEVH 304  
DB 595 DNLEIQFQSDNLTKDSSSEISSTGAFLSNQTSSEASSNSMGSINSPLSLSLTNSESA 654  
QY 305 GNAEVHASFPDGGVSAGFSNSNSTVAIDHSLSL--AGERTWAEWTGLNTADTARLANA 362  
DB 655 TNQ-----SNSSEATKVDNNSST-----HSSNILNSGNSDSDSDSDSNL--699  
QY 363 NIRVNTGTAPIYVNLPTTSLVGNQTLATIKADENQLSQ-----ILAPNYPYKSLAP 418  
DB 700 -----SSSPNLETNQIISSEVSENNISENPKVSSNSVQENSTDEH 742



Query Match	4.7%;	Score 179;	DB 2;	Length 1635;
Best Local Similarity	19.4%;	Pred. No. 0.44;		
Matches 176;	Conservative 137;	Mismatches 315;	Indels 278;	Gaps 44;
QY	8	LLNESESSQGLGYFFSDLNFPQAPWVTSSTTGDLSTPSSLENIPISENQYFQSAIWSG	67	
DB	510	LTAENKITTQGHKASAGGDWIDA-----NDVKIGVQKTSNQETTDGKHENUGLG	560	
QY	68	FTKVKKSDBYETAT-----SADNHVTMWVDDQEVIN-----KASNSNKIRLEKGRLYQIKI	118	
DB	561	GYDHNNDKYAEKTSHSSEITADGNILISVKDDVAITGSKVKATGDFVQAKEG---GIKI	617	
QY	119	QYQRENPTKEK-----GLDFKLYWTDSQNK-----KEVISSDNQLPBLKOKSNS	163	
DB	618	DNAISTTTNKVDERTGTGVAFDI--TGSSKANNSEKSTGSEVVSANLKI--ISKDVDV	673	
QY	164	RKKRSTAG-----PIVPD-----RDNGIPDS-----LEVEGTV--DVNKRFTFLSPWISN	209	
DB	674	IGSLVKAGELGIETLGDINVAAGQEKQIKDEQKTLITDGTSDGKNQ-----	723	
QY	210	IHEKKGL--TKYKSSPEKWTASDPYSDPEKVTGRI--DKNVSPEARHPILVAAYPIVHDM	266	
DB	724	--YQAGLKVEHTSESEKTEKVINHGSTLEGGIVKLEADKDVFTGSG--LNTTKGDADITA	780	
QY	267	ENILISKNDQSQNTD-----SETRTSKNTSTSTERTHTS	301	
DB	781	ENVSPFAAODTTTTSNKEKFTGVNHAHTGTGMDKAGSCAGVSVEETKIDSEKSTAVVSHT--	839	

Search completed: May 3, 2004, 19:41:16  
Job time : 16.0855 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 19:27:02 ; Search time 8.2539 Seconds

(without alignments)  
4636.784 Million cell updates/sec

Title: US-09-848-909A-10

Perfect score: 3774

Sequence: 1 EVKQENRLNSESSESSQGL.....TSTNGIKLILFSSKKGYEIG 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3758	99.6	764	1 PAG_BACAN	P13423 bacillus an
2	235.5	6.2	204	1 YPBI_BACAN	P13422 bacillus an
3	185.5	4.9	1630	1 MSP1_PLAFK	P04932 plasmodium
4	185.5	4.9	1639	1 MSP1_PLAFW	P04933 plasmodium
5	174	4.6	1658	1 YMF7_YEAST	Q03661 saccharomyc
6	169.5	4.5	1230	1 SMC3_YEAST	P47037 saccharomyc
7	166	4.4	1000	1 S155_YEAST	Q07833 saccharomyc
8	165.5	4.4	2334	1 WAPA_BACSU	P07833 bacillus eu
9	164	4.3	1803	1 YUL3_YEAST	P47024 saccharomyc
10	162.5	4.3	2116	1 MYS2_DICDI	P08799 dictyosteli
11	162	4.3	1957	1 SPOF_SCHPO	Q10411 schizosacch
12	161.5	4.3	1979	1 P115_MYCHR	P41508 mycoplasma
13	160	4.2	1276	1 BXD_CLOBO	P19321 clostridium
14	159.5	4.2	2869	1 RBPI_PLAVB	Q00798 plasmodium
15	159	4.2	1385	1 PAT1_SCHPO	O13735 schizosacch
16	158.5	4.2	1167	1 CAGA_HELPU	Q921t1 helicobacte
17	157	4.2	2867	1 RBP2_PLAVB	Q00799 plasmodium
18	155.5	4.1	1024	1 RIP3_MOUSE	P97434 mus musculu
19	155	4.1	1790	1 USOI_YEAST	P25386 saccharomyc
20	154.5	4.1	1208	1 PCPI_SCHPO	Q92351 schizosacch
21	154.5	4.1	1460	1 MDS3_YEAST	P53094 saccharomyc
22	154.5	4.1	1487	1 KCC4_YEAST	P25389 saccharomyc
23	154	4.1	1037	1 YF16_YEAST	P43359 saccharomyc
24	153.5	4.1	1233	1 YF16_YEAST	Q9ere6 rattus norv
25	153	4.1	1039	1 RIP3_RAT	O67838 aquifex aeo
26	152.5	4.0	1116	1 YK54_AQUAE	P38931 saccharomyc
27	152	4.0	1420	1 SRB9_YEAST	P38931 saccharomyc
28	150	4.0	918	1 YMJ6_CABEL	P34487 caenorhabdi
29	150	4.0	1882	1 Y468_MYCPN	P75109 mycoplasma
30	150	4.0	1928	1 MYS1_YEAST	P08964 saccharomyc
31	149.5	4.0	678	1 INCY_YEAST	P53968 saccharomyc
32	149	3.9	1250	1 BXE_CLOBO	Q00496 clostridium
33	149	3.9	1577	1 HLYA_PROMI	P16466 proteus mir

#### ALIGNMENTS

##### RESULT 1

PAG\_BACAN STANDARD; PRT; 764 AA.  
AC P13423: Q9F5R7; Q9KH69; Q9ROU2;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Protective antigen precursor (PA) (PA-83) (PA83) (Anthrax toxins  
DE translocating protein) [Contains: PA-20 (PA20); PA-63 (PA63)].  
GN PAGA OR PAG OR PXOI-110.  
OS Bacillus anthracis.  
OC Plasmid PXOI.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99172073; PubMed=3148491;  
RA Welkos S.L., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H.,  
RA Schmidt J.J.;  
RT "Sequence and analysis of the DNA encoding protective antigen of  
RT Bacillus anthracis.";  
RL Gene 69:287-300(1988).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=28\_33, BA1024, and BA1035;  
RX MEDLINE=99214082; PubMed=10197996;  
RA Price L.B., Hugh-Jones M., Jackson P.J., Keim P.;  
RT "Genetic diversity in the protective antigen gene of Bacillus  
RT anthracis.";  
RL J. Bacteriol. 181:2358-2362(1999).  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V770-NP1-R / ATCC 14185;  
RX MEDLINE=20359347; PubMed=10899854;  
RA Cohen S., Mendelson I., Alboum Z., Kobiler D., Elhanany E., Bino T.,  
RA Leitner M., Inbar I., Rosenberg H., Gozes Y., Barak R., Fisher M.,  
RA Kronman C., Velan B., Shaffer A.;  
RT "Attenuated nontoxigenic and nonencapsulated recombinant Bacillus  
RT anthracis spore vaccines protect against anthrax.";  
RL Infect. Immun. 68:4549-4558(2000).  
[4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sterne;  
RX MEDLINE=99445483; PubMed=10515943;  
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,  
RA Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,  
RA Martinez Y., Ricks D., Svensson R., Jackson P.J.;  
RT "Sequence and organization of pXOI, the large Bacillus anthracis  
RT plasmid harboring the Anthrax toxin genes.";  
RL J. Bacteriol. 181:6509-6515(1999).  
[5]  
RP DOMAINS.  
RX MEDLINE=91332080; PubMed=1651334;  
RA Singh Y., Kimpel K.R., Quinn C.P., Chaudhary V.K., Leppla S.H.;  
RT "The carboxy-terminal end of protective antigen is required for  
RT receptor binding and anthrax toxin activity.";

Q96173 homo sapien  
P40468 saccharomyc  
Q06366 clostridium  
P47551 mycoplasma  
P11978 saccharomyc  
P50495 plasmodium  
Q9ZK45 helicobacte  
P18494 saccharomyc  
Q9CF20 lactococcus  
Q03213 saccharomyc  
Q53654 staphylococ  
P58302 thermoplasm

RL J. Biol. Chem. 266:15493-15497(1991).  
 [6] CHARACTERIZATION.  
 RP STRAIN=Sterne;  
 RX MEDLINE=94327640; PubMed=8051159;  
 RL Milne J.C., Furlong D., Hanna P.C., Wall J.S., Collier R.J.;  
 RT "Anthrax protective antigen forms oligomers during intoxication of  
 RL mammalian cells";  
 J. Biol. Chem. 269:20607-20612(1994).  
 [7] CHARACTERIZATION.  
 RP MEDLINE=2112592; PubMed=11207581;  
 RX Beauregard K.E., Collier R.J., Swanson J.A.;  
 RL "Proteolytic activation of receptor-bound anthrax protective antigen  
 RT on macrophages promotes its internalization";  
 Cell. Microbiol. 2:251-258(2000).  
 [8] TOXIN REGULATION.  
 RP STRAIN=Weybridge;  
 RX Koehler T.M., Dai Z., Kaufman-Yarbray M.;  
 RL "Regulation of the Bacillus anthracis protective antigen gene: CO2 and  
 RT a trans-acting element activate transcription from one of two  
 RT promoters";  
 J. Bacteriol. 176:586-595(1994).  
 [9] MUTAGENESIS OF PHE-342; PHE-343 AND ASP-344.  
 RP STRAIN=Sterne;  
 RX MEDLINE=95050722; PubMed=7961869;  
 RL Singh Y., Klimpel K.R., Arora N., Sharma M., Leppla S.H.;  
 RT "The chymotrypsin-sensitive site, FFD315, in anthrax toxin protective  
 RT antigen is required for translocation of lethal factor";  
 J. Biol. Chem. 269:29039-29046(1994).  
 [10] MUTAGENESIS OF DOMAIN 4 LOOPS.  
 RP STRAIN=Sterne;  
 RX MEDLINE=99185012; PubMed=10085028;  
 RL Varughese M., Teixeira A.V., Liu S., Leppla S.H.;  
 RT "Identification of a receptor-binding region within domain 4 of the  
 RT protective antigen component of anthrax toxin";  
 Infect. Immun. 67:1860-1865(1999).  
 [11] MUTAGENESIS OF TRP-375; MET-379 AND LEU-381.  
 RP STRAIN=Sterne;  
 RX MEDLINE=2102804; PubMed=11178978;  
 RL Batra S., Gupta P., Chauhan V., Singh A., Bhatnagar R.;  
 RT "Trp 346 and Leu 352 residues in protective antigen are required for  
 RT the expression of anthrax lethal toxin activity";  
 Biochem. Biophys. Res. Commun. 281:186-192(2001).  
 [12] MUTAGENESIS OF PHE-581; PHE-583; ILE-591; LEU-595 AND ILE-603.  
 RP STRAIN=Sterne;  
 RX MEDLINE=21438996; PubMed=11554763;  
 RL Ahuja N., Kumar P., Bhatnagar R.;  
 RT "Hydrophobic residues Phe552, Phe554, Ile562, Leu566, and Ile574 are  
 RT required for oligomerization of anthrax protective antigen";  
 Biochem. Biophys. Res. Commun. 287:542-549(2001).  
 [13] MUTAGENESIS OF PRO-289.  
 RP STRAIN=Sterne;  
 RX MEDLINE=21255689; PubMed=11356563;  
 RL Khanna H., Chopra A.P., Arora N., Chaudhry A., Singh Y.;  
 RT "Role of residues constituting the 2beta1 strand of domain II in the  
 RT biological activity of anthrax protective antigen";  
 FEBS Microbiol. Lett. 199:27-31(2001).  
 [14] MUTAGENESIS OF GLN-512; ASP-541; LEU-543 AND ARG-621.  
 RP MEDLINE=21125576; PubMed=1122612;  
 RX Mogridge J., Mouriz M., Collier R.J.;  
 RT "Involvement of domain 3 in oligomerization by the protective antigen  
 RT moiety of anthrax toxin";  
 J. Bacteriol. 183:2111-2116(2001).  
 [15]

RP MUTAGENESIS OF LYS-426; ASP-454 AND PHE-456.  
 RX MEDLINE=21269403; PubMed=11113126;  
 RA Sellman B.R., Nassi S., Collier R.J.;  
 RT "Point mutations in anthrax protective antigen that block  
 RT translocation";  
 J. Biol. Chem. 276:8371-8376(2001).  
 [16] MUTAGENESIS OF PRO-213; PHE-216; PHE-231; LEU-232; PRO-234; ILE-236;  
 RP ILE-239; TRP-255 AND LEU-265.  
 RX STRAIN=Sterne;  
 RL MEDLINE=22112896; PubMed=12117959;  
 RA Chauhan V., Bhatnagar R.;  
 RT "Identification of amino acid residues of anthrax protective antigen  
 RT involved in binding with lethal factor";  
 Infect. Immun. 70:4477-4484(2002).  
 [17] X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
 RP MEDLINE=97132099; PubMed=9039918;  
 RA Petosa C., Collier R.J., Klimpel K.R., Leppla S.H., Liddington R.C.;  
 RT "Crystal structure of the anthrax toxin protective antigen";  
 Nature 385:833-838(1997).  
 [18] REVIEW.  
 RP MEDLINE=21428689; PubMed=11544370;  
 RX Mock M., Pouet A.;  
 RT "Anthrax";  
 Annu. Rev. Microbiol. 55:647-671(2001).  
 CC -!- FUNCTION: One of the three proteins composing the anthrax toxin,  
 CC the agent which infects many mammalian species and that may cause  
 CC death. PA binds to a receptor (ATR) in sensitive eukaryotic  
 CC cells, thereby facilitating the translocation of the enzymatic  
 CC toxin components, edema factor and lethal factor, across the  
 CC target cell membrane. PA associated with LF causes death when  
 CC injected, PA associated with EF produces edema. PA induces  
 CC immunity to infection with anthrax.  
 CC -!- SUBUNIT: Anthrax toxins are composed of three distinct proteins, a  
 CC protective antigen (PA), a lethal factor (LF) and an edema factor  
 CC (EF). None of these is toxic by itself. PA+LF forms the lethal  
 CC toxin (LeTx); PA+EF forms the edema toxin (EdTx). PA-63 forms  
 CC heptamers and this oligomerization is required for LF or EF  
 CC binding. Once activated, at low pH, the heptamer undergoes  
 CC conformational changes and converts from prepore to pore inserted  
 CC in the membrane, forming cation-selective channels.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- DOMAIN: The molecule is folded into four functional domains. Each  
 CC domain is required for a particular step in the toxicity process.  
 CC Domain 1 contains two calcium ions and the proteolytic activation  
 CC site. Cleavage of the PA monomer releases the subdomain 1a, which  
 CC is the N-terminal fragment of 20-kDa (PA20). The subdomain 1b is  
 CC part of the remaining 63-kDa fragment (PA63) and contains the  
 CC binding sites for LF and EF. Domain 2 is a beta-barrel core  
 CC containing a large flexible loop that has been implicated in  
 CC membrane insertion and pore formation. There is a chymotrypsin  
 CC cleavage site in this loop that is required for toxicity. Domain 3  
 CC has a hydrophobic patch thought to be involved in protein-protein  
 CC interactions. Domain 4 appears to be a separate domain and shows  
 CC limited contact with the other three domains: it would swing out  
 CC of the way during membrane insertion. It is required for binding  
 CC to the receptor; the small loop is involved in receptor  
 CC recognition.  
 CC -!- PTM: Proteolytic activation by furin or a furin-like protease  
 CC cleaves the protein in two parts, PA-20 and PA-63; the latter is  
 CC the mature protein. The cleavage occurs at the cell surface and  
 CC probably in the serum of infected animals as well; both native and  
 CC cleaved PA are able to bind to the cell receptor. The release of  
 CC PA20 from the remaining receptor-bound PA63 exposes the binding  
 CC site for EF and LF, and promotes oligomerization and  
 CC internalization of the protein.  
 CC -!- MISCELLANEOUS: In Ref.9 multiple mutagenesis experiments were  
 CC performed that showed that the residues present in the small loop  
 CC of domain 4, and not the ones in the large loop, are involved in  
 CC receptor recognition.  
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL BINARY TOXIN FAMILY.

Query Match 99.6%; Score 3758; DB 1; Length 764;  
Best Local Similarity 99.6%; Pred. No. 1.2e-177;  
Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVKQENLLNESBSSOGLGYFSDLNQFQAPMVVVTSTTGDLSIPSELENIPSENOYF 60  
DB 30 EVKQENLLNESBSSOGLGYFSDLNQFQAPMVVVTSTTGDLSIPSELENIPSENOYF 89

QY 61 QSAIWSGFIKVKSDYFATFADNHVMTWYDDQEVINKASNSKIRLEKRLYQIKIY 120  
DB 90 QSAIWSGFIKVKSDYFATFADNHVMTWYDDQEVINKASNSKIRLEKRLYQIKIY 149

QY 121 QRENPTKGLDFKLYWTDSONKKEVISSDNILQELPKOKSSNRKSTSGAGPTVPDRN 180  
DB 150 QRENPTKGLDFKLYWTDSONKKEVISSDNILQELPKOKSSNRKSTSGAGPTVPDRN 209

QY 181 DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 240  
DB 210 DGIPDSLEVEGYVDVKNKRTFLSPWISNIHEKGLTKYKSSPEKWTASDPYDFEKVT 269

QY 241 GRIDKNVSPERHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRISKNVTSRHT 300  
DB 270 GRIDKNVSPERHPLVAAPIVHVDMENIILSKNEDQSTQNTDSETRISKNVTSRHT 329

QY 301 SEVHGNAEVHASFDDIGSVSAGFNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 360  
DB 330 SEVHGNAEVHASFDDIGSVSAGFNSNSTVAIDHLSLAGERTWAETMGLNTADTARL 389

QY 361 NANIRYVNTGTAPIYVNLPTSLVLGKQNTLATIKADENQLSQILAPNNYYPSSKILAPIA 420  
DB 390 NANIRYVNTGTAPIYVNLPTSLVLGKQNTLATIKADENQLSQILAPNNYYPSSKILAPIA 449

QY 421 LNAQKDFSTPTIMYNOFLEKTKQLRLDQVYGNIAVYNTENGVRVDTGNSNVEV 480  
DB 450 LNAQKDFSTPTIMYNOFLEKTKQLRLDQVYGNIAVYNTENGVRVDTGNSNVEV 509

QY 481 LPQIOETTARIIFNGKDLNVERIAAVNPSDPLETTKPDMTLKEALKIAGFNEPENG 540  
DB 510 LPQIOETTARIIFNGKDLNVERIAAVNPSDPLETTKPDMTLKEALKIAGFNEPENG 569

QY 541 QYQKDIITFDFNFQOQTSQNTKQALAEANVYVLDKIKNAKMLIRDKRHFYDR 600  
DB 570 QYQKDIITFDFNFQOQTSQNTKQALAEANVYVLDKIKNAKMLIRDKRHFYDR 629

QY 601 NNIAVGADESVMKEAHRVINSSTGLLNIIDKIRKILSGYIVIEDTEGLKEVINDRY 660  
DB 630 NNIAVGADESVMKEAHRVINSSTGLLNIIDKIRKILSGYIVIEDTEGLKEVINDRY 689

QY 661 DMLNSSLRQDGTFFDKYNDKPLIYSNPNKYVAVTKENTIIINPSENGDTSTNG 720  
DB 690 DMLNSSLRQDGTFFDKYNDKPLIYSNPNKYVAVTKENTIIINPSENGDTSTNG 749

QY 721 IKKILIFSKKGVEIG 735  
DB 750 IKKILIFSKKGVEIG 764

RESULT 2  
ID YPBI\_BACAN STANDARD; PRT; 204 AA.  
AC P13422; Q9X377.  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Hypothetical protein pXOI-111.  
GN pXOI-111.  
OS Bacillus anthracis.  
OC Plasmid pXOI.  
CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1392;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=89172073; PubMed=3148491;  
RA Welkos S.J., Lowe J.R., Eden-McCutchan F., Vodkin M., Leppla S.H.,  
Schmidt J.J.;  
RT "Sequence and analysis of the DNA encoding protective antigen of  
Bacillus anthracis.";  
RL Gene 69:287-300(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sterner;  
RX MEDLINE=99445483; PubMed=10515943;  
RA Okinaka R.T., Cloud K., Hampton O., Hoffmaster A.R., Hill K.K.,  
Keim P., Koehler T.M., Lamke G., Kumano S., Mahillon J., Manter D.,  
Martinez Y., Ricke D., Svensson R., Jackson P.J.;  
RT "Sequence and organization of pXOI, the large Bacillus anthracis  
plasmid harboring the Anthrax toxin genes.";  
RL J. Bacteriol. 181:6509-6515(1999).  
CC -1- SUBCELLULAR LOCATION: Membrane-bound (Potential).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@sib-sib.ch).  
CC EMBL; M22589; AAA22636.1; --  
DR EMBL; AF065404; AAD32415.1; --  
DR PIR; G59104; G59104.  
DR PIR; I39933; I39933.  
DR HSP; P13423; IACC.  
KW Hypothetical protein; Plasmid; Transmembrane.  
FT TRANSMEM 162 182 POTENTIAL.  
FT CONFLICT 93 93 T -> A (IN REF. 1).  
FT CONFLICT 184 204 KSCNCILYVEVSQMLNSVFY -> NHVIVYLSM (IN  
REF. 1).  
SQ SEQUENCE 204 AA; 23029 MW; E1657B23AB4273FD CRC64;  
Query Match 6.2%; Score 235.5; DB 1; Length 204;  
Best Local Similarity 34.6%; Pred. No. 2.6e-05;  
Matches 56; Conservative 36; Mismatches 47; Indels 23; Gaps 6;

QY 587 MNTILIRKRFHYDNNIAVGADESVMKEAHRVINSSTGLLNIIDKIRKILSGYIVIE 646  
DB 1 MNTILIRKRFHYDNNIAVGADESVMKEAHRVINSSTGLLNIIDKIRKILSGYIVIE 59

QY 647 EDTE-----GLKEVINDRYDMLNSSLRQDGTFFDKYNDKPLIYSNPN 693  
DB 60 KPSNHLTNSPVITITLAGKDSGVGELYRVLS-----DGTGFLDFNKFDPENRSLV-DPG 112

QY 694 YKNNYVAVTKEN-TIINPSENGDTSTNGIKILIFSKKGVEI 734  
DB 113 DDVYVAVTKEDFNAVTRDENGNI-A-NKLNVLVLGKIKEI 153

RESULT 3  
ID MSP1\_PLAFK STANDARD; PRT; 1630 AA.  
AC P04932;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)  
DE (P190).  
GN MSP-1.  
OS Plasmodium falciparum (isolate K1 / Thailand).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86136024; PubMed=3004972;  
RA Mackay M., Goman M., Bone N., Hyde J.E., Scaife J., Certa U.,  
Stunnenberg H., Bujard H.;

RT "Polymorphism of the precursor for the major surface antigens of  
RT Plasmodium falciparum merozoites: studies at the genetic level.";  
RL EMO J. 4:3823-3829(1985).  
RN [2]  
RP REVISIONS, SEQUENCE FROM N.A.  
RA Pan W., Toile R., Bujard H.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
CC (Potential).  
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42  
CC kDa and 19 kDa antigens which are the major surface antigens of  
CC merozoites. The maturation take place during schizont.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X03371; CAA27070.1; -  
DR InterPro: IPR006209; EGF\_like.  
DR Pfam: PF00008; EGF; 1.  
DR Malaria; Merozoite; Glycoprotein; Repeat; Signal; Glycoprotein;  
KW Transmembrane; GPI-anchor.  
FT SIGNAL 1 19  
FT CHAIN 20 1630  
FT DOMAIN 67 84  
FT TRANSMEM 1614 1630  
FT CARBOHYD 97 97  
FT CARBOHYD 259 259  
FT CARBOHYD 755 755  
FT CARBOHYD 759 759  
FT CARBOHYD 774 774  
FT CARBOHYD 835 835  
FT CARBOHYD 911 911  
FT CARBOHYD 955 955  
FT CARBOHYD 1049 1049  
FT CARBOHYD 1156 1156  
FT CARBOHYD 1165 1165  
FT CARBOHYD 1436 1436  
FT CARBOHYD 1517 1517  
SQ SEQUENCE 1630 AA; 187289 MW; ADBDEC3CE0A46322 CRC64;  
Query Match 4.98; Score 185.5; DB 1; Length 1630;  
Best Local Similarity 20.5%; Pred. No. 0.12; Indels 265; Gaps 47;  
Matches 170; Conservative 138; Mismatches 256;  
5 ENRL-LNESSSSQGLLYGYSDFDLNFOAPMVVTSSTTGLSIPSSSELENIPSENYQFQSA 63  
918 ENLISLGNKNIYQELIGQKSSSE-NF-----YKILKDSDTFYNE 956  
64 IWSGFIKVKSDYFATPSADNHNVTWVDDOEVINKASNS-----NKIRLEKGRLY- 114  
957 SFTNFVKSADD-----INSLNDSKRRKLEEDINKLTLQLSFLYKYLKLERLFD 1011  
115 -----QKIQYQENPTEKGLDFKLYWTDSONKKEVISSDNLQPLKQKSSNSRKKRS 168  
1012 KKTIVGVYQVQIKLTLKQLESKL--NSLNPKHVL--QNSVFPNKKKEAIEAETEN 1067  
169 TSAGTVPDRNDGIPDSLEVEGVTVVKNKRTPLSPWISNIHEKKGLTKY----KSP-- 223  
1068 T-----LENTKILKHY-----KGLVKKYNGESSPLK 1094  
224 ---EKWSTASDPYDFE-----KVTGRIDQNVSPPEAR-----HPLVAAYPIVHVD 265  
1095 TLSEESTQEDENYASLENFKVLKLEGLKDNLNLEKKLSYLSGLHLHIA-----E 1147  
266 MENILSKNEQDQNTDSETRTSKNTSTSRHTSEVHNAEVAHSFFDIGSVSAGFS 325  
1148 LKEVIKKNK--YTGNSPSENNT-----DVNNALESYKKFLPEGTDTVATWS 1191

QY 326 NNSSTVAIDHSLSLAGERTWAETWMTGLNTPADTARLNANIRYVNTGTAPIYVLPPT---- 381  
DB 1192 ESGSDTLEQSQPKKPASTHVGAE--NTITTSQ-NVDDEVDDVIIVPIFGSEEDYDDL 1247  
QY 382 -SLVLGNQTLATIKADENQLSQILAPNPNYPSKNLAPIA--LNAQKDFSSTPTIMYN 437  
DB 1248 GQVVTGEAVTPSVI---DNILSKI---ENEYEVLYLKLPLAGVYRSLSKQLENNVMTFVN 1301  
QY 438 -----QFLELEKTKQLRLDQV-YGNIAI-----YNFENGGRVVDGSGNWSVL 481  
DB 1302 VKDILNSRFNKNFKNV-LESGLIPYKOLTSNYSVVKDPYKFLNKEKDKFLLSSYNTIK 1360  
QY 482 PQIQTETARIIFNGKLNILVERIAAVNPSDPL-----ETTKPDM-TLKALKAIAGP 533  
DB 1361 DSD-----TDINFA-----NDVLGYKILSEKYSKSLDSIKYINDKQGE 1401  
QY 534 NEPN-----GNLOYQKDKITE-----FDNFQDQTSQNIKNQLAELNATNIYT 576  
DB 1402 NEKYLPLANNIETLYKTVDKIDLFVHLEAKVLYNYTEK---SNVEVKIKELN--YLKT 1456  
QY 577 VLDKIKLNAQON--ILIRKRFHYDRNN-----IAGV-ADSVVKEAHEVINSSTEGLL 628  
DB 1457 IQDKLADFKNNNFVGIADLSTDYNNHLLTKFLSTGMVFENLAKTVLSNLLDGNLQGM 1515  
QY 629 LNIQDKD--IRKILSGVIVETEDTEGLKEVINDRYDMLN--SSLFQDQKTFIDFKYNDKLP 686  
DB 1516 LNIQHQCVK-----QCPQNSGCFRHLDE-----REECKLLNYKQEGDKC- 1557  
QY 687 LYSNPNYKYNVAVKENTIIINPSNG-----DTSTNGIKKI 724  
DB 1558 --VENPNPTCN-----EN-----NGGCDADAKCTEEDSGSNG-KKI 1590  
RESULT 4  
MSPL PLAFW STANDARD; PRT; 1639 AA.  
ID MSP1\_PLAFW AC P04933;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Merozoite surface protein 1 precursor (Merozoite surface antigens)  
DE (PMMSA) (P195).  
GN MSP-1.  
OS Plasmodium falciparum (isolate Wellcome).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5848;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86014355; PubMed=2995820;  
RA Holder A.A., Lockyer M.J., Odink K.G., Sandhu J.S., Riveros-Moreno V.,  
RA Nicholls S.C., Hillman Y., Davey L.S., Tizard M.L.V., Schwarz R.T.,  
RA Freeman R.R.;  
RT "Primary structure of the precursor to the three major surface  
RT antigens of Plasmodium falciparum merozoites.";  
RL Nature 317:270-273(1985).  
RN [2]  
RP REVISIONS.  
RA Holder A.A.;  
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor  
CC (Potential).  
CC -!- PTM: Merozoite surface antigen contain the sequence of 83 kDa, 42  
CC kDa and 19 kDa antigens which are the major surface antigens of  
CC merozoites. The maturation take place during schizont.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```
DR EMBL; X02919; CAA26676.1; --
DR PIR; A24594; A24594.
DR PIR; S05603; S05603.
DR PDB; 1CEJ; 28-MAY-99.
DR Pfan; PFO0008; EGF; 1.
DR InterPro; IPR006209; EGF_like.
DR Malaria; Merzoite; Polyprotein; Repeat; Signal; Glycoprotein;
KW Transmembrane; GPI-anchor; 3D-structure.
KW SIGNAL
FT SIGNAL 1 19
FT CHAIN 20 1639
FT CARBOHYD 116 116
FT CARBOHYD 268 268
FT CARBOHYD 764 764
FT CARBOHYD 768 768
FT CARBOHYD 783 783
FT CARBOHYD 844 844
FT CARBOHYD 920 920
FT CARBOHYD 964 964
FT CARBOHYD 1058 1058
FT CARBOHYD 1165 1165
FT CARBOHYD 1174 1174
FT CARBOHYD 1445 1445
FT CARBOHYD 1526 1526
FT CARBOHYD 1526 1526
SQ SEQUENCE 1639 AA; 187618 MW; 2C25SB6616C87F6E CRC64;

Query Match
Best Local Similarity 4.9%; Score 185.5; DB 1; Length 1639;
Matches 170; Conservative 138; Mismatches 256; Indels 265; Gaps 47;

QY 5 ENRL-LNESSSQGLLYFSDLPQAPMVVTSSTGDLSPSELENIPEENQYFOSA 63
DB 927 ENILSGKNKNIYQELIGOKSSE-NF-----YEKILKDSPTFYNE 965
QY 64 IWSGFTKVKSDSYTATGADNHVTWVDDQEVINKASNS-----NKILEKRLV- 114
DB 966 SFTNFVKSADD-----INSLNDSKRRKLEIDINKLKTQLSFLYKYLKLERLPD 1020
QY 115 -----QIKIQYQRENPTEKGLDFKLYWTDSONKEVSSDNLQPELKQKSSNRKRS 168
DB 1021 KKKTGVKYNQIKKLTLLKEQESKL--NSLNPKHVL--QNFSPFFNKKKEAEIETEN 1076
QY 169 TSAGPTVPDRNDGIPDSLEVEGYTVGVNKNKTFSPWISNIHEKKGLTKY---KSSP-- 223
DB 1077 T-----LENTKILLKHY-----KGLVKYNGSSPLK 1103
QY 224 ---EKMTASDPYSDFE-----KVTGRIDKNVSPAR-----HPLVAAYPIVHD 265
DB 1104 TLSEESIQTEDNYASLENPKVLSKLEGKLDNLEKKKLSVLSGLHLIA-----E 1156
QY 266 MENILSKNQDQSTQNTDSETRISKNTSTSRHTSEVHGNAEVHASPFDIGSVSAGFS 325
DB 1157 LKEVIKKN---YTGNSPSENNT-----DVNNALESYKKFLPEGTVDVATVVS 1200
QY 326 NSNSSTVAIDHSLSLAGERWAEITMGITADTARLNANIRYVNTGTAPIYVLPPT--- 381
DB 1201 ESGSDTLEQSQPKKPASTHVGAES---NITTSQ-NVDDEVDDVIIPIFGSESEYDDL 1256
QY 382 -SLVLGKNGQTLATIKADENOLSOILAPNNYPKSNLAPIA---LNQKQPSPTPTMNVN 437
DB 1257 GOVVTGEAVTFSVI---DNILSKI---ENEYEVLYLKPLAGVYRSILKQLENNVMFTFN 1310
QY 438 -----QFLELEKTKQLRDLTDQV-YGNIAI-----YNFENGVRVDTGNSWSEVL 481
DB 1311 VKDILNRFKRENFKNV-LESLLIPYKDLTSSNYVVKPYKLNKEKDKFLSSNYIK 1369
QY 482 PQIQETTARIIFNGKDANLVERRIAANVPSDPL-----ETTKPDM-TLKEALKIATOF 533
DB 1370 DSID-----TDINFA-----NDVLGYVKILSEKYKSDLSIKKYINDKQGE 1410
QY 534 NEPN---GNLYOGKDKITE-----FDNFDOQTQNTKNQNLAEINATNIYT 576
DB 1411 NEKVPFLNNIETLYKTVNDKIDLFIHLHEAKVLNITYEK---SNVEVKIKELN---YLKT 1465
```

```

Db 460 EQQENDEPEKDDIIRSLDKNFHGNKSEYSENVLNETDPAIVERNO-IND---VE 515
Qy 191 GYTVGVKQKRTFL---SPWISNIHE--KGLTKYKSS-----PEKWSFASD----- 231
Db 516 GIDVTGKSVESDLHEHSP--DNLYDLARAMLQFOQSRSNCPQKEEQVSSYLGHNGS 573
Qy 232 -----PYSDFEKTGTGRDKNVSPEARHPLVAAYPIVHVDMENIILSKNEDQ 277
Db 574 NLSGRSLDSEBQIPLKDF--TGNNNNLKTDRGDLSSS---VEIEVEKV---SEKK 622
Qy 278 STQNTDSETRTISKNTSRTSHTSVGHNAEVAHAFDGGSVSAGFSNNSSTVADHS 337
Db 623 LGSTEKELVPLSTDTTN-----NSSLNDSIYSLDD- 657
Qy 338 LSLAGERTWAETMGLNTADTARLANIRYVNTGTAPIYVLTFTSLVLGKQNTLATIKAD 397
Db 658 -----ADAISENLT-----VPLMEIKTPKYEVISSV-----YSSTSYE 694
Qy 398 ENQLS-----QILAP-----NNYPSKN-----LAPALNAQKDFSTPTTM 434
Db 695 DNTVAMPPOVEVTSPPMNDPFNSLNDYDKHDLKSLTALALAPFTKKDAEFVEAGVTK 754
Qy 435 -----NNQFLELEKTKQLRLDQVYGNIAVYFNGRVRVDTGNNMSEVLPTQIE 486
Db 755 SCITSTSGHTNIFHTSKETQVYS-DLDESTENVTFENETG---DENKNQSKNPPGVAN 809
Qy 487 TTARIIPNGKDLNLVERRIAANVPSDPLETTKPDMLKEALKIAGFNEPKNLOYQKGD 546
Db 810 STDKSTEDNTD---EKYFSAINYN---VTGDSCEDIIETAGNVEE---NLRYCEKD 858
Qy 547 ITFFDNF--DOOTSQ---IKQLA-----ELNATNIYTVLDIKLNKMNILIRDKR 595
Db 859 MNEAEMSSGDECVKQNDGSKTKIISFSTDSPDNFQESNDNTFFSTK----- 905
Qy 596 FHYDRNNAIAGDESVMKAHR-EVIN-----SSTEGLLINIDKD 634
Db 906 --YKVRNSDLEDESXKELTKAEVVDKLBDEESDSEYQDYADPEPQNGDESNIVKG 963
Qy 635 IRKILSGYIEIBDTTGLKEVINDRYDMLNLSRLQDGTDFDKYKNDKLPYTSNPNY 694
Db 964 TKK-----DTLGIPEPEKYN-----KVHEEELTFEANYSS 995
Qy 695 KVNVAVTKENTIN 709
Db 996 SVNQNKMDHTDVIN 1010

```

## RESULT 6

```

SMC3_YEAST STANDARD; PRT; 1230 AA.
ID SMC3_YEAST
AC P47037;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Structural maintenance of chromosome 3 (DA-box protein SMC3).
GN SMC3 OR YJL074C OR J1049.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RX MEDLINE=97474309; PubMed=9335333;
RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RA Michaelis C., Ciosk R., Nasmyth K.;
RT "Cohesins: chromosomal proteins that prevent premature separation of
  sister chromatids.";
RL Cell 91:35-45(1997).
RN [2].
RP SEQUENCE FROM N.A.
RX MEDLINE=96208490; PubMed=8641269;
RA Galibert F., Alexandraki D., Baur A., Boles E., Chalwatzis N.,

```

```

RA Chuat J.-C., Coster F., Cziepluch C., De Haan M., Domdey H.,
RA Durand P., Enkian K.-D., Gatus M., Goffeau A., Grivell L.A.,
RA Hennemann A., Herbert C.J., Heumann K., Hilger F., Hollenberg C.P.,
RA Huang M.-E., Jacq C., Jauniaux J.-C., Katsoulou C., Kirchath L.,
RA Kleine K., Kordes E., Koetter P., Liebl S., Louis E.J., Manus V.,
RA Kewes H.-W., Miesga T., Obermaier B., Perez J., Pohl T.M.,
RA Portetelle D., Pujol A., Purnelle B., Ramezani Rad M., Raamussen S.W.,
RA Rose M., Rossau R., Schaaff-Gerstenschlaeger I., Smits P.H.M.,
RA Scarcez T., Soriano N., Tovan D., Tzeremia M., Van Broekhoven A.,
RA Vanderbol M., Wedler D.H., Von Wettstein D., Wambutt R., Zagulski M.,
RA Zollner A., Karpfinger-Hartl L.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
  X.";
RL EMBL J. 15:2031-2049 (1996).
RN [3].
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1678;
RA Sor F.J.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4].
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC1; IRR1 AND MCD1, AND
  INTERACTION OF THE COHESIN COMPLEX WITH SCC2.
RX MEDLINE=99145468; PubMed=990856;
RA Toth A., Ciosk R., Uhlmann F., Galova M., Schleiffer A., Nasmyth K.;
RT "Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to
  establish cohesion between sister chromatids during DNA replication.";
RL Genes Dev. 13:320-333(1999).
RN [5].
RP IDENTIFICATION IN A COHESIN COMPLEX WITH SMC3; MCD1 AND IRR1, AND
  STRUCTURE.
RX MEDLINE=21980168; PubMed=11993169;
RA Haering C.H., Leewe J., Hochwagen A., Nasmyth K.;
RT "Molecular architecture of SMC proteins and the yeast cohesin
  complex.";
RL Mol. Cell 9:773-788(2002).
CC -!- FUNCTION: Involved in chromosome cohesion during cell cycle and in
  DNA repair. Central component of cohesin complex. The cohesin
  complex is required for the cohesion of sister chromatids after
  DNA replication. The cohesin complex apparently forms a large
  proteinaceous ring within which sister chromatids can be trapped.
  At anaphase, the complex is cleaved and dissociates from
  chromatin, allowing sister chromatids to segregate.
CC -!- SUBUNIT: Cohesin complexes are composed of the SMC1 and SMC3
  heterodimer attached via their hinge domain. MCD1/SCC1 which link
  them, and IRR1/SCC3, which interacts with MCD1. The cohesin
  complex also interacts with SCC2, which is required for its
  association with chromosomes.
CC -!- SUBCELLULAR LOCATION: Nuclear protein. Associates with chromatin.
  Before prophase it is scattered along chromosome arms. At
  anaphase, the MCD1 subunit of the cohesin complex is cleaved,
  leading to the dissociation of the complex from chromosomes,
  allowing chromosome separation.
CC -!- DOMAIN: The flexible hinge domain, which separates the large
  intramolecular coiled coil regions, allows the heterotypic
  interaction with the corresponding domain of SMC1, forming a V-
  shaped heterodimer. The two heads of the heterodimer are then
  connected by different ends of the cleavable MCD1 protein, forming
  a ring structure (By similarity).
CC -!- SIMILARITY: Belongs to the SMC family. SMC3 subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
RL EMBL; Y14278; CAA74655.1; -.
DR EMBL; Z49349; CAA89366.1; -.
DR EMBL; X88851; CAA61313.1; -.
DR PIR; S56850; S56850.
DR GeneOnline; 141688; -.

```

```
DR SGD; S0003610; SMC3.
DR GO; GO:0007151; P:sporulation (sensu Saccharomyces); IMP.
DR GO; GO:0007130; P:synaptonemal complex formation; IMP.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR003405; SMC_C.
DR InterPro; IPR003395; SMC_N.
DR Pfam; PF02483; SMC_C; 1.
DR Pfam; PF02463; SMC_N; 1.
DR KW Mitosis; Cell cycle; Chromosome partition; ATP-binding; Coiled coil;
KW Nuclear protein.
FT NP BIND 32 39 ATP (POTENTIAL).
FT DOMAIN 172 482 COILED COIL (POTENTIAL).
FT DOMAIN 453 684 FLEXIBLE HINGE.
FT DOMAIN 655 1041 COILED COIL (POTENTIAL).
FT DOMAIN 1126 1161 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 1230 AA; 141336 MW; B152D88F7780341F CRC64;

Query Match 4.5%; Score 169.5; DB 1; Length 1230;
Best Local Similarity 19.8%; Pred. No. 0.51;
Matches 182; Conservative 136; Mismatches 355; Indels 247; Gaps 38;

QY 1 EVKQENRLNSESSESSQGLGYFDLNFQPMVVT--SSTGDLSPSSLEINIPENQ 58
DQ 209 EMEQKELEKYNELERNKIQFTLYDLRELNEVINQMERLDGYN-----NTVYSSEQ 262
QY 59 YFQ-----SAIWGFIKVKSDYETPATSADNHVMTWVDQEVINK--- 99
DQ 263 VIQELDKREDIMDQVSKLSSI-EASLAKIKATDLOQAKLRESEISQKLTNNVNIKDVQ 321
QY 100 ---ASNNKIRLEKGLQYIK-IQYQRENPTKGLDFKLYWTDSONKEVSSDNLQLOPE 155
DQ 322 QQIESNEEQRNDSDATLKEIKSIIEQRKQKSLKILPRYOELT-----KEEAMV--KLQLAS 375
QY 156 LAKSSNRKKESTGAGTPVPRDNDGIPDSLEVEGYTVDKNKRTFLSPWISNHE-KK 214
DQ 376 LQOKQDLILKGEYARKFKSDERTWIHSEIE-----ELKSS-----IQNLNELS 422
QY 215 GUTKYSSPEKSTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAPIVHVDMENIILSKN 274
DQ 423 QLQMDRTSLRKQYSAID-----EIEELIDSLNGPDTKGQL-----EDFDELHLKQK 471
QY 275 EQQSTQNTDSERTLSKNTSRTTSEVHGNAEVHASFDDIGGSVAGFNSNSSTVAI 334
DQ 472 LSESLDTRKELWKEQKLTQVLETLLSDVNQNR-----NVNETMSRSLANGIINVKBI 525
QY 335 DHSLSLAGE-----RTWAETWG-----LNTADTARLANIRYVNTG 370
DQ 526 TEKLKISPSVFGTIGELIKYNDKYKTCAEVIGGNSLPHIVVDTEETATLIMNELYRMKG 585
QY 371 TAPIYVNLPTTSLVKGKQNTATIKADENQSLAPNNYPSKVLAPIALNAQKQFSST 430
DQ 586 GRVTF--IFPLNRLSDSDVKFSSNTTQTQFTPLIKIKYEPFEKA-----VKHVFQKT 638
QY 431 PITMNYNQFLEKTKQLR---LDTDOV-YGNIAF--YNFENGRVVRVDTGNWSVLPQI 484
DQ 639 IVVKDLGGQGLAKKHKLNAILDGDGRADKRGVLTGGVLDQHKTRLESKLNLESRSQH 698
QY 485 QETTARIIFNGKDLNLVERRIAAVNPS-----DPLETTKPD-MPLK 524
DQ 699 KKILBELDVRNELNDIDTKIDQVNGNIRKVSNDRESVLTNIEVYRTSLNTKQNEKLTLE 758
QY 525 EALK-IAFGNPNGLQYQKDIETEDFN---EDQQTSONIKQLA-----E 568
DQ 759 ESLNAILKLEKLNTRTFAQKLNFTFNDLLQDFDSELSKEERLESSTKE-SAAHNK 818
QY 569 LNATN-----IYTVLDKIKLNKAMNILIRDKRFHYDRNNIAYG-----A 607
DQ 819 LNITSALBGIITTD--SLNAELESKLIPQENDLESKMSEVGDAFIFGLQDELKELQLE 876
QY 608 DRSVYKEAREVINSST-----EGLLL-----NIDKIRKI 638
DQ 877 KESVEKQENAVLELGTVQREIESIAETNNKLEKANNQORLLKLLDNFQKSVKKT 936
```

```
QY 639 LSGYIIVEIEDTEGLKE-----VINDRYDM-----LN-----ISSLRQDGK 673
DQ 937 MIKKTTLVTRREELQRIREIGLIPEDALVNDPFSDDITSQQLQLRNDMATEISGLKNVVK 996
QY 674 -TFIDFKYNDK-----LPLYSNPNYKNNVYAVTVKNIINPSENG 714
DQ 997 RAFENFKFNERRKDLAERASELDESKDSIQDLVLVKLKKQKYNVAV-----DSTFQKVSFNP 1052
QY 715 DT-----STNGIKKILIFSK 729
DQ 1053 EAVFERLVRPGTAKLIHRK 1072

RESULT 7
SI55 YEAST
ID SI55 YEAST STANDARD; PRT; 1000 AA.
AC P43612;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE SIF4-associated protein SAPI55.
GN SAPI55 OR YPFO40W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96220458; PubMed=8649382;
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Arndt K.T.;
RT "The S4P, a new family of proteins, associate and function positively
RL Mol. Cell. Biol. 16:2744-2755(1996).
RN [2]
RP SEQUENCE OF 98-1000 FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yanazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae.";
RL Nat. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 98-1000 FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=96287652; PubMed=8686379;
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of
RL chromosome VI from Saccharomyces cerevisiae.";
RL Yeast 12:149-167(1996).
CC -!- FUNCTION: Associates with the SIF4 phosphatase in a cell cycle
CC dependent manner. May be directly or indirectly involved in
CC SIF4-dependent functions in budding and in normal G1 cyclin
CC expression.
CC -!- PTM: Hyperphosphorylated in the absence of SIF4.
CC -!- SIMILARITY: Belongs to the SAPS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U50560; RAC49303.1; --
CC EMBL; D50617; BAA09279.1; --
CC GenOnline; 140194; --
CC SGD; S0001936; SAPI55.
CC InterPro; IPR007587; SAPS.
CC -----
```



DR Pfam: PF04499; SAPs; 1.  
KW Phosphorylation; Cell cycle.

FT CONFLICT 663  
FT CONFLICT 674 689  
FT CONFLICT 814 815  
FT CONFLICT 818 822  
SQ SEQUENCE 1000 AA; 114880 MW; AD4BF0BC7C598D5 CRC64;

Query Match

Best Local Similarity 19.5%; Pred. No. 0.57;

Matches 173; Conservative 128; Mismatches 323; Indels 262; Gaps 42;

QY 1 EVKQENLLNESSSQGLGYFSS-DL-----NFOAPMVVTSSTGTD-----LSIPS 47  
DB 90 EKEEERSNSESSTTSFSSGSTSKTDLDEEDISNATAPMVVTKNLDNFIERMLVET 149  
QY 48 SELENIPSNQYFQSAINGSFIVKYSDEYTFATSADNHTVWVDD-QEVINKASNSKI 106  
DB 150 ELNELSRQKTLDFICGFFDFKTKNKV-----NNMEYLVQLMCEISKIKTATTV 203  
QY 107 RLEKGLYQIKYQRENTEKGLDFKLYWTDSONKKEVSSDNLQELPQKSNRKK 166  
DB 204 DLNN-----LIDYEQOQLD-----DSSQEDVYVESDTQEKEEDNNNNKK 247  
QY 167 R-----STGAGTVPDRDNDGIPDSLEVEGYTVDVNKRRTFLSP-----WISNIHERKG 215  
DB 248 RKRGSSTSGNDINNDDDDANEDDESAYL---TKATIIEIFSLDWLIS-----ES 299  
QY 216 LKYKSSPK-KSTASDPYSDFEKTVGRIDKXVSEARHPLVAAPVIVHVDMENTILSKN 274  
DB 300 LVKQSYLNKIWSIINQP-----NFNSE-NSPLV---PIFLKINQNLTLR- 341  
QY 275 EDQ-----STQWTDSET-----STQWTDSET-----RTISKNT 293  
DB 342 QOYLNFIFERSFVDDMLKHVDISLLMDFFLKIISTOKIESPTGIELVYDQNLISKCL 401  
QY 294 STSRHTSEVHGAENVHAFDFIGSVSAGFNSSTVAL-DHSLG--LAGERTWATM 350  
DB 402 SFLNNKESP-----ADIQACVGDPLKALIAISANAPLDDISIGFNSLTRQLASPEIAKLV 457  
QY 351 GLNTADTARNIRVYVGTAPIYVNLPTLSVLGKQNTLATIKADENQLSOILAPNNY 410  
DB 458 DI-----MINQGAALNT-----TVSIVI---ELIRKNSDQVQNLITTTIKT 498  
QY 411 YPSKNLAPIALNAQKDFSTPTIMYVNFLEKTKQLBLDQDVYGNIAVTFENGVR 470  
DB 499 HPPSNRDPYLVGLLRKFNHLSDFPQIILDIENDANIPLENQHLHEKFKPLGFERFKV- 557  
QY 471 VDTGSNWEVLPOIQTETARIIFNGKDLNVERRIAANVP--SDPLETTKPDWTL--KEA 526  
DB 558 -----VELIAELLHCSNGLMNSKRAERIAARRDKVRSQLSHLQDALNLSIEKEQ 610  
QY 527 LKIAFG-----FNEPKNLQYQKDI-----TEPFDNFD---QOTSONIKNQLAE 568  
DB 611 LKTKHSPTRDTHLKNNGKIDNNDNDDSDYDGEIDSEFEPYINMKQNIKLRTPD 670  
QY 569 LNATNIYVLDKILAKNKNILRDKRFYDRNNAVAGD-----ESVYKEAHEVIN-- 621  
DB 671 TVGT-----YSKSNYMRD---CFQNNELFTHPWNFWNFWNFDIIQIQFNQR 716  
QY 622 ---SSTEGLLNID--KDIRKILSGYVIEIE---DTEGLKEVINDRYDMLNLS---LR- 669  
DB 717 MDSYNSFLVLSFLNLSKSYQFQVTDIVISEKGTVDVRSFPIVDNDFDKITDFILRG 776  
QY 670 QGKTFIDPKKNY-----DKLPLYISN-----PNYKNVY 699  
DB 777 YQDSYKPYELRKNMGLYMGHIVLIAEVEWFKSLYKVDYISRYSSNLSQTEWQYVSEVL 836  
QY 700 AVTK-----ENTIIINPSENGDSTN 719  
DB 837 NERPMYSKILGGSYIDGNGNLIIPQLPNTTIVLP--NGDASN 880

## RESULT 8

WAPA\_BACSU STANDARD; PRT; 2334 AA.  
ID WAPA\_BACSU  
AC Q07833;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Wall-associated protein precursor.  
GN WAPA OR N17G OR BSU39230.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
CX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=93302506; PubMed=8316082;  
RA Foster S.J.;  
RT "Molecular analysis of three major wall-associated proteins of  
RT Bacillus subtilis 168; evidence for processing of the product of a  
RT gene encoding a 258 kDa precursor two-domain ligand-binding  
RT protein.";  
RL Mol. Microbiol. 8:299-310(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / BGSC1A1;  
RX MEDLINE=95219088; PubMed=7704263;  
RA Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;  
RT "Cloning and sequencing of a 29 kb region of the Bacillus subtilis  
RT genome containing the hut and wapa loci.";  
RL Microbiology 141:337-343(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / BGSC1A1;  
RX MEDLINE=97124196; PubMed=8969509;  
RA Yoshida K.-I., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,  
RA Miwa Y., Fujita Y.;  
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome  
RT containing the lic and cel loci, and creation of a 177 kb contig  
RT covering the gnt-sacXY region.";  
RL Microbiology 142:3113-3123(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Enrian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Guiseppe G., Guy B.J., Haga K., Haele J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holstappel S., Hosono S., Rullo M.F., Itaya M., Jones L.,  
RA Joris B., Kazamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
RA Medina N., Melillo R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin M., Taconi E., Takagi T., Takanashi H., Takemaru K.,  
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vanderbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengruber T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

"The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.";  
 RL Nature 390:249-256(1997).  
 CC -!- FUNCTION: STILL UNKNOWN. NOT INVOLVED IN CELL MEMBRANE METABOLISM,  
 CC SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MAY BE RELEASED  
 CC INTO THE MEDIUM.  
 CC -!- DOMAIN: HAS TWO LIGAND-BINDING DOMAINS; THE N-TERMINUS, HAS THREE  
 CC 101 AA REPEATS WHICH ARE RESPONSIBLE FOR CELL WALL BINDING; THE  
 CC C-TERMINUS CONSISTS OF TWO BLOCKS OF RESIDUES WITH A CONSERVED  
 CC MOTIF REPEATED 31 TIMES.  
 CC -!- SIMILARITY: THE REPEATED MOTIF IN THE C-TERMINUS HAS SOME  
 CC SIMILARITY TO THE REPEAT IN *E. COLI* Rhs GROUP OF PROTEINS (RhsA-D).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; L05634; AAA22883.1; -  
 CC EMBL; D31856; BAA06656.1; -  
 CC EMBL; D29885; BAA06260.1; -  
 CC EMBL; D83026; BAA11683.1; -  
 CC EMBL; Z99124; CAB15959.1; -  
 CC PIR; S32920; S32920. -  
 CC Subtilisin; BG10797; wapa.  
 CC InterPro; IPR003305; CBM\_CenC.  
 CC InterPro; IPR006530; YD.  
 CC Pfam; PF02018; CBM 4\_9; 1.  
 CC Pfam; PF05593; Rhs repeat; 14.  
 CC TIGRfams; TIGR01643; YD repeat; 2x; 17.  
 CC Cell wall; Repeat; Signal; Complete proteome.  
 CC OR 32 (POTENTIAL)  
 CC SIGNAL 1 28  
 CC CHAIN 29 2334 WALL-ASSOCIATED PROTEIN.  
 CC DOMAIN 504 869 3 X 101 AA APPROXIMATE TANDEM REPEATS.  
 CC REPEAT 504 869 1-1.  
 CC REPEAT 605 869 1-2.  
 CC REPEAT 636 736 1-3.  
 CC REPEAT 769 869 1-3.  
 CC DOMAIN 1021 2139 31 X 21 AA APPROXIMATE TANDEM REPEATS OF  
 CC X(4)-G-X(4)-(YF)-X-D-X(2)-G-X(4).  
 CC  
 CC 2-1.  
 CC REPEAT 1021 1040  
 CC REPEAT 1042 1061  
 CC REPEAT 1063 1082  
 CC REPEAT 1083 1102  
 CC REPEAT 1109 1128  
 CC REPEAT 1129 1148  
 CC REPEAT 1150 1169  
 CC REPEAT 1174 1193  
 CC REPEAT 1199 1218  
 CC REPEAT 1219 1238  
 CC REPEAT 1646 1665  
 CC REPEAT 1667 1686  
 CC REPEAT 1690 1709  
 CC REPEAT 1711 1730  
 CC REPEAT 1732 1751  
 CC REPEAT 1753 1772  
 CC REPEAT 1795 1814  
 CC REPEAT 1820 1839  
 CC REPEAT 1840 1859  
 CC REPEAT 1861 1880  
 CC REPEAT 1887 1906  
 CC REPEAT 1908 1927  
 CC REPEAT 1929 1948  
 CC REPEAT 1969 1982  
 CC REPEAT 1983 2002  
 CC REPEAT 2008 2027  
 CC REPEAT 2028 2047  
 CC REPEAT 2051 2070  
 CC REPEAT 2071 2090  
 CC REPEAT 2093 2112  
 CC 2-2.  
 CC 2-3.  
 CC 2-4.  
 CC 2-5.  
 CC 2-6.  
 CC 2-7.  
 CC 2-8.  
 CC 2-9.  
 CC 2-10.  
 CC 2-11.  
 CC 2-12.  
 CC 2-13.  
 CC 2-14.  
 CC 2-15.  
 CC 2-16.  
 CC 2-17.  
 CC 2-18.  
 CC 2-19.  
 CC 2-20.  
 CC 2-21.  
 CC 2-22.  
 CC 2-23.  
 CC 2-24 (APPROXIMATE).  
 CC 2-25.  
 CC 2-26.  
 CC 2-27.  
 CC 2-28.  
 CC 2-29.  
 CC 2-30.

FT REPEAT 2120 2139 2-31.  
 SQ SEQUENCE 2334 AA; 258329 MW; B75138CCDD278BAA3 CRC64;  
 Query Match 4.4%; Score 165.5; DB 1; Length 2334;  
 Best Local Similarity 18.2%; Pred. No. 1.9;  
 Matches 173; Conservative 126; Mismatches 336; Indels 317; Gaps 42;  
 Qy 1 EVKQENRLNSESSESSQGLLYGYPFDLNFQA-----PMVTSSTTGDLSIPSELENIP 54  
 Db 269 EVERSCKVSKLEKNEBEGYLLHLTADENMLKDPERYVPSIDPSTLSVSSDTFVMSAYP 328  
 Qy 55 SENQYFQSAIWSGFIK--VKKSDEYTFATSAADNHVTM----- 89  
 Db 329 TTNYSASSQKWDANLKAAYLVKGTGYD-KTGTNYAFMKFNKLKPIQNMVTVKATLKTYVA 387  
 Qy 90 -----WYDDEVINK-----ASNSKIKLEGR--LYQIKIQYOR 122  
 Db 388 HSYVGTGKATGLWD--TVNSYDYNKAVTWTNPKASNGIKADVHGQWASVDVTAARKS 444  
 Qy 123 ENPTEKGLDFKLYWTDSONK---KEVISSDNLQPELKOKSSNRKRSSTAGTVP--- 176  
 Db 445 WNSGANYGFKLH-TNGNGKEYWKLIIS-----SANSANKPYIEVTYTPKGN 491  
 Qy 177 -----DRDNDGIPP-----SLB-VEGYTVDVNKRFTLSPMISNHEKKGLTKYKSSP 223  
 Db 492 TPTIKAYHNGDSTGYFDISWKVEGAKGYKWTYNGKEYQAIAGNV-----TSWSTKG 545  
 Qy 224 EK-WSTASDPYSDPEKVTGRID-KNVSPEARHPLVAAYPIVHVDMENIILSKNE----- 275  
 Db 546 KKIWPTSAEIAASKRYKL--HLDGDKGAELALDP-----SPVYKNSGGSYATSKYNIWGVA 599  
 Qy 276 --DQSTONTDSETRTISKNTSTRTSTSEVHGNAEVAHAFDIDGGVSAG-----PS 325  
 Db 600 IFDQEGAMSAKAPKVPINPVGKAQAPSAKYNNGNA-TGYFDLSKAVSGATQYKQVFN 658  
 Qy 326 NSNSSTVAIDH--SLSLAGERTWAETMGLNTADTA-----RLNANIRYVNTGTA 372  
 Db 659 KGFTETDLGNGQISWTGKKIWIPTSAEIKAGKVALHLKDGSAEPIFNGPYKAGGD 718  
 Qy 373 PIYNVLTSLVLGRNQTALIKADENQLSQILAPNN-----YPSK-----NLA--- 417  
 Db 719 GAKNYSFKIIAYNKGGEALASPAATPALDIAKPNVTGYLTNTKSSOTGYVNLWEK 778  
 Qy 418 -----PIALNAQKDESSPTI-----TMVYNQFLELEKTK-----QLRL 450  
 Db 779 VQNAKGYKNIYNGKEYQSFVGDADHWTTONKNIMPTSEIKAGSYKLTGDKGBELAL 838  
 Qy 451 DTDOVYGN-----IATYN--FENGVRVDTGSNWSEV-L 481  
 Db 839 DSPVYNNANGYKKNYSFTLVAYDANGETIPTAFPFTFHEGAEFLGTETYSIIDI 898  
 Qy 482 P--QIQTETARIIPNGKDLNL-----VERRIAIVNPSDPL-----ETT----- 517  
 Db 899 PSGQLNGATGNVI VNEEDLSIDGRGPGGLSRTYNSLDSDHLPFGQGWYADAETSVISTD 958  
 Qy 518 -----KPDMTLKEALKIAPG-FNEPNGNLQYQOKDITEFDNFDDQTSQNIKNQLAEL- 569  
 Db 959 GGAMVIDEDATTHRTKADGTQYPTGVILELTETADQFILTKDQTNAYFNKKGKQLQ 1018  
 Qy 570 -----NATNIYTVLDKIKLNKAKNIIIRDKRHHYDRN-----NIAVGADESV 611  
 Db 1019 KVDGHNNAVTYTYNDKNQLTAITDASGRKLTFYTDENGHVTSITGPKNKVTVSYENDL 1078  
 Qy 612 VKE-----AHR--EVINSSTEGILL 629  
 Db 1079 LKVTDDTGTVSYDYDSEGLVQYQSANSTEAKPEVTEYQYSGHREKALNACKETVYV 1138  
 Qy 630 NIDKDI-----EKILSGY-----IVEIEDTEGLKEVINDRYDMLNI 665  
 Db 1139 SYDADKKTLLMTQPNGRKQVGYNEAGNPQIIVDADGLKITTNTKYEKNV 1190

RESULT 9

YJL3 YEAST  
ID YJL3 YEAST STANDARD; PRT; 1803 AA.  
AC P47024; P87192;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Transposon Ty4 207.7 kDa hypothetical protein.  
GN TY4B OR YJL113W OR J0780.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288c / FY1679;  
RX MEDLINE=97103775; PubMed=8948101;  
RA Czepluch C., Kordes E., Pujol A., Jauniaux J.-C.;  
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X  
RT reveals 19 open reading frames including URA2 (5' end), TRK1, PBS2,  
RT SPT10, GCD14, RPE1, PHO86, NCA3, ASF1, CCT7, GZF3, two tRNA genes,  
RT three remnant delta elements and a Ty4 transposon.";  
RL Yeast 12:1471-1474 (1996).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z49389; CAA89409.1; .  
DR SGD; S0003649; YJL113W.  
DR InterPro; IPR001584; Rve.  
DR InterPro; IPR001878; Znf\_CCHC.  
DR Pfam; PF00665; rve; 1.  
DR SMART; SM00343; Znf\_C2HC; 1.  
KW Transposable element; Hypothetical protein.  
SQ SEQUENCE 1803 AA; 207693 MW; 16DCD7284A8D52D3 CRC64;  
  
Query Match 4.3%; Score 164; DB 1; Length 1803;  
Best Local Similarity 19.3%; Pred. No. 1.6; Indels 270; Gaps 41;  
Matches 162; Conservative 123; Mismatches 274;  
Qy 1 EVKQENRLNESSSQGLGYFSDLNFOAPMVVTSSTGDLSPSSLENIPENQVF 60  
Db 1078 QLKXTN--HETSPKESGIG--TNVGRNTNNEI-SLKTGTSLSPIKTESINHH--- 1128  
Qy 61 QSALWGFIVKKSDEYTFATSDNHNVTWVDQEVINKASNSKIRLEKRLYQIKQY 120  
Db 1129 -----SNDY-----STNKVE-----KF 1140  
Qy 121 QREN-----PTEKGLDFLYWTDSONKKEVSSDNLQLPKQKSSNRKSRSTSAGTYP 176  
Db 1141 EKENHHPPPIEDIVDM-----SDQTMESNCQGNLKLKLVTDKNVPTDNGTNSPR-- 1193  
Qy 177 DRDNDGTPDSLEVEGYTVVKNKTFELSPWISNIHEKK-----GLTKY----- 219  
Db 1194 -----LEONIEASGSPVQTVNKSFAFLNKFSSLNMMKREKRDKNNSLTSYELERDKR 1247  
Qy 220 -----KSSPEKWTASDPYSDPEKVTGRIDKNVSPPEARHPLVAAPYIVHVDNENILSK 273  
Db 1248 SKNRVKLIPDNMETVSAKIRAIYNEAISKNPDLKEHKEYQAY---HKEQLNKLDMK 1304  
Qy 274 NEDOSTONTSETTISKN--TSTSRHTSEVHGNAEVHASFDDIGSVSAGSNSST 331  
Db 1305 VFVDVVKYSRSE---IPDNLIVPTNTFTTKRNG---IYKARIVCRGDTQSPDYTVIT 1358  
Qy 332 VAIDHS-----LSLAGERT-WAETMGLNTADT-ARLNANIRYNTGTAPYVNLPL-TTSL 383  
Db 1359 ESLNHNHKKIFELMTANNRNMFKTLDINHAFLYAKLEEE-----IYIHPHRRRC 1408  
Qy 384 VLGRNQTALIKADENQ-----LSQI-LAPNNYYP-----SKNLAPIALNAQKDFS 428

Db 1409 VVKLNKALYGLKQSPKEWHDHRLQVINGLGLKDNSTYTPGLYQTEDEKNLM-IAVYVDDCVI 1467  
Qy 429 STPTITMNYNQL-ELEKTKQLR-----LDTOVYGNIAATYFNGRVR----- 469  
Db 1468 AASNEQRDLDEFINKLSNFELKITGLTIDVLDTD-ILGMDLVYKRLGTIDTLTKSPFN 1526  
Qy 470 RVDTSNWNSEVLPOIQTETARIIFNGKOLNVERRIAANP-SDPLETTKPD-----MTL 523  
Db 1527 RMD--KXNEELKKRKSPHSMSTYK-----IDPKKDVLOESEEPRQGVKL 1573  
Qy 524 KEALKIAGFNBPNGNLQYQGGKITEFDNFPOQTSQNIKNQLAELNATNYTVLDKIKL 583  
Db 1574 QQLL-----GELNVV-RHKCRYDIEFAVKKVARLVNYPHERVFYMKIIOYL-- 1620  
Qy 584 NAKWNLIRDKRPHYDRN-----NIAVGADSVKAEHRE-----VI-----NS 622  
Db 1621 -----VRYKDIGHYDRDNCKKKVIAITDASVGYDAQSGRIGVILWTGMIFNVYSNK 1675  
Qy 623 STEGLLNIDKDIRKILSGV-----IVEIEDT-----EGLK----- 653  
Db 1676 STNRCVSSTEAEHLAIVEGVADSETLKVTLKELGEGDNDIVMTDSKPAIQCLNRSYQ 1735  
Qy 654 -----EVINDRYDMLNLISSLRQDGKTF-----DEKKY 681  
Db 1736 PKKFTWIKTEIIKEKIKESIMLLKITGKNIAIDLTPVSAASDFRF 1784  
  
RESULT 10  
MYS2\_DICDI STANDARD; PRT; 2116 AA.  
ID MYS2\_DICDI  
AC P08799;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Myosin II heavy chain, non muscle.  
GN MHCa.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87092266; PubMed=3540939;  
RA Warrick H.M., de Lozanne A., Leinward L.A., Spudich J.A.;  
RT "Conserved protein domains in a myosin heavy chain gene from  
RT Dictyostelium discoideum.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437 (1986).  
RN [2]  
RP PHOSPHORYLATION SITES, AND MUTAGENESIS.  
RC STRAIN=AX2;  
RX MEDLINE=90353583; PubMed=2387408;  
RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,  
RA Gerisch G.;  
RT "Replacement of threonine residues by serine and alanine in a  
RT phosphorylatable heavy chain fragment of Dictyostelium myosin II.";  
RL FEBS Lett. 269:239-243 (1990).  
RN [3]  
RP PHOSPHORYLATION SITES.  
RX MEDLINE=88112226; PubMed=2828113;  
RA Wagie G., Noegel A., Scheel J., Gerisch G.;  
RT "Phosphorylation of threonine residues on cloned fragments of the  
RT Dictyostelium myosin heavy chain.";  
RL FEBS Lett. 227:71-75 (1988).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.  
RX MEDLINE=93345066; PubMed=7619795;  
RA Fisher A.J., Smith C.A., Rhoden J.B., Smith R., Sutcho K., Holden H.M.,  
RA Rayment I.;  
RT "X-ray structures of the myosin motor domain of Dictyostelium  
RT discoideum complexed with MgADP.Befx and MgADP.AIF4-.";  
RL Biochemistry 34:8960-8972 (1995).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.

RX MEDLINE=95345067; PubMed=7619796;  
 RA Smith C.A., Rayment I.;  
 RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the  
 truncated head of Dictyostelium discoideum myosin to 2.7-A  
 resolution.";  
 RT Biochemistry 34:8973-8981(1995).  
 RL [6]  
 RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
 RP MEDLINE=96206189; PubMed=86115130;  
 RA Smith C.A., Rayment I.;  
 RT "X-ray structure of the magnesium(II). ADP.vanadate complex of the  
 Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";  
 RL Biochemistry 35:5404-5417(1996).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
 RX MEDLINE=97452860; PubMed=9305951;  
 RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;  
 RT "X-ray crystal structure and solution fluorescence characterization  
 of the Dictyostelium discoideum myosin motor domain.";  
 RL J. Mol. Biol. 274:394-407(1997).  
 CC -!- FUNCTION: Myosin is a protein that binds to actin and has ATPase  
 activity that is activated by actin.  
 CC -!- SUBUNIT: Myosin II heavy chain is two-headed. It self-assembles  
 into filaments. Hexamer of 2 heavy chain subunits (MHC), 2 alkali  
 light chain subunits (MLC) and 2 regulatory light chain subunits  
 (MLC-2).  
 CC -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL  
 CC CORTX.  
 CC -!- DOMAIN: Each myosin heavy chain can be split into 1 light  
 meromyosin (LMM) and 1 heavy meromyosin (HMM). It can be further  
 split into 2 globular subfragments (S1) and 1 rod-shaped  
 subfragment (S2).  
 CC -!- DOMAIN: The rodlike tail sequence is highly repetitive, showing  
 cycles of a 28-residue repeat pattern composed of 4 heptapeptides,  
 characteristic for alpha-helical coiled coils.  
 CC -!- PTM: Phosphorylation inhibits thick filament formation and reduces  
 the actin-activated ATPase activity.  
 CC -!- MISCELLANEOUS: Dictyostelium myosin II has no K(2)EDTA ATPase  
 activity, perhaps correlated with the absence of a Cys at the SH-1  
 position (688).  
 CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.  
 CC -!- SIMILARITY: Contains 1 IQ domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M14628; AAA33227.1; -;  
 CC PIR; A26655; A26655.  
 DR PDB; 1MMA; 03-DEC-97.  
 DR PDB; 1MMD; 17-AUG-96.  
 DR PDB; 1MNG; 03-DEC-97.  
 DR PDB; 1MMN; 03-DEC-97.  
 DR PDB; 1MND; 17-AUG-96.  
 DR PDB; 1MNE; 23-DEC-96.  
 DR PDB; 1VOW; 17-AUG-96.  
 DR PDB; 1LXK; 28-JAN-98.  
 DR PDB; 1DOK; 20-DEC-00.  
 DR PDB; 1DOY; 20-DEC-00.  
 DR PDB; 1DOZ; 20-DEC-00.  
 DR PDB; 1D1A; 20-DEC-00.  
 DR PDB; 1D1B; 20-DEC-00.  
 DR PDB; 1D1C; 20-DEC-00.  
 DR PDB; 1FMV; 20-DEC-00.  
 DR PDB; 1FMW; 20-DEC-00.  
 DR PDB; 1G8X; 17-JAN-01.  
 DR PDB; 1JMY; 07-NOV-01.  
 DR PDB; 1JX2; 07-NOV-01.  
 DR DictyBase; DDB0002015; mbca.  
 DR InterPro; IPR000048; IQ\_region.  
 DR InterPro; IPR001609; myosin\_head.  
 DR InterPro; IPR004009; Myosin\_N.  
 DR InterPro; IPR008989; Myosin\_S1\_N.  
 DR Pfam; PF00612; IQ; 2.  
 DR Pfam; PF00663; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR PRODOM; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 DR PROSITE; Coiled coil; Actin-binding; ATP-binding; 3D-structure;  
 Myosin; Methylation; Phosphorylation.  
 KW Calmodulin-binding; Methylation; MYOSIN HEAD-LIKE.  
 FT DOMAIN 1 761  
 FT DOMAIN 762 791  
 FT DOMAIN 817 2116  
 FT NP\_BIND 179 186  
 FT DOMAIN 660 660  
 FT DOMAIN 738 752  
 FT MOD\_RES 130 130  
 FT MOD\_RES 1823 1823  
 FT MOD\_RES 1833 1833  
 FT MOD\_RES 2029 2029  
 FT TURN 3 5  
 FT TURN 7 8  
 FT HELIX 10 15  
 FT HELIX 25 28  
 FT STRAND 34 37  
 FT TURN 41 42  
 FT STRAND 48 55  
 FT STRAND 59 63  
 FT STRAND 69 73  
 FT HELIX 74 76  
 FT STRAND 78 79  
 FT HELIX 83 85  
 FT TURN 86 87  
 FT STRAND 90 90  
 FT HELIX 91 93  
 FT HELIX 99 110  
 FT TURN 111 113  
 FT STRAND 116 119  
 FT TURN 120 121  
 FT STRAND 122 126  
 FT HELIX 137 142  
 FT TURN 143 145  
 FT HELIX 148 150  
 FT HELIX 155 169  
 FT TURN 170 170  
 FT STRAND 173 179  
 FT TURN 181 182  
 FT HELIX 185 200  
 FT HELIX 210 226  
 FT STRAND 227 228  
 FT TURN 234 235  
 FT STRAND 236 237  
 FT STRAND 240 247  
 FT TURN 249 250  
 FT STRAND 253 261  
 FT HELIX 265 268  
 FT TURN 269 269  
 FT TURN 274 274  
 FT STRAND 278 278  
 FT HELIX 279 287  
 FT HELIX 290 296

```

FT TURN 297 297
FT HELIX 301 303
FT TURN 305 307
FT TURN 316 317
FT HELIX 320 334
FT TURN 335 335
FT HELIX 338 355
FT TURN 356 356
FT STRAND 360 360
FT STRAND 368 368
FT HELIX 373 382

Query Match 4.3%; Score 162.5; DB 1; Length 2116;
Best Local Similarity 19.4%; Pred. No. 2.3;
Matches 138; Conservative 113; Mismatches 219; Indels 243; Gaps 33;

QY 86 HVTWVDDOEVINKASNSKIRLEKGLYQIKIYOQRENTPEKGLDFKLYWTDSONKKE- 144
DB HVNQLSEEEKQKESNEKRVKVDLEK-EVSELKQIEEVASKAV-----TEANKKES 1346
QY 145 -----VSSDNLQIPELKQKSSNRKRTSAGTVPDRNDGIPDSLEVEGYT 193
DB 1347 ELDEIKRQYADVSSRDKSVLEQLTKLQAKNEELRNTA-----BEAQGLDRAERS--- 1396
QY 194 VDVGNKRTF-LSPWISNHE-----KKGLTKYKSSPEKWSASDPVSPFVKYT 240
DB 1397 ---KKAEFDLEBAVKNLBEETAKVKYKAKMKAETDYRSKSELDDAKNVSSQYQVI 1453
QY 241 GRIDKNVS-----PEARHPLVAAPIVHVDMNIIILSKNEDOSTONTSETRTTSK 291
DB 1454 KRLNEELSELRSVLSEADERCNSAIKAKKTAESALESL---KDEIDAAANAKAKERSK 1510
QY 292 NTS*SRHTSHVGHNAEVAHAFDDIGSVSAGFSNSNSTVAIDHLSLAGERTWAETWG 351
DB 1511 EL-----EVR-VASELESLEDKSGTVNVEIRKDD-----AEIDD 1544
QY 352 LNTADTARLANIRVYNTGTAPIYVLPITSLVLGKNOQLATIKADENGLSQILAPNYY 411
DB 1545 LR-----ARLD-----RETESRIKSEDEK----- 1563
QY 412 PSKNLAFIALNAQDFS--STPITMNYNQFLEKTKQLRLDQVYGNIAIYNFENGRV 469
DB 1564 -----KNTRQKQFADLEAKVEEAQREVTTIDRLKX-KLESDDI--DLST----- 1603
QY 470 RVDTGSNWSEVLPQIQTARI-IFNGK---DNLVVERIAAVNPDPLETKFDMTLKE 525
DB 1604 QLDT-----ETKGRIKIEKSKKLEQTLAERRAAEEGSSKAAD-----EE 1643
QY 526 ALKIAFGFNEPNGNLQYQGDITDFDNFDQ-----TSQNIKNQLAELNATNIYTVLD 579
DB 1644 IRKQVW-----QEVDELRAQLDSERAAALNASEKKIKSLVAE-----VD 1681
QY 580 KIKLNKANIILIRKRFHYDRN---NTAVGADESIVKEAHREVINSSTGLLLNIDKQIR 636
DB 1682 EVKQLEDEILAKOKLVAKRAULEVELEVRDQLEEDSRSELEDSKRLTTEVE-DIK 1740
QY 637 KILSGYIVIEI-----DTEGLKEVINDRYDMLNTSS-----LRQDGKTFI- 676
DB 1741 K---KYDAEVEQNTKLEAKKCLTDVDTLTKQLEDEKKLNSEKAKKLESENEDFLA 1797
QY 677 -----DFKYNKDLPLYISNPNYKVNVTAKNTIINPSENGD 715
DB 1798 KLDAEVKNRSAEKDRKKYKNDL-----KDTKYKLNDEAATKQTQTEIGAARLE 1846

RESULT 11
SPOF SCHPO
ID SPOF SCHPO STANDARD; PRT: 1957 AA.
AC Q10411_Q3US89;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sporulation-specific protein 15.

```

```

GN OS SPO15 OR SPAC1F3.06C
OC Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RX MEDLINE=20107136; PubMed=10639340;
RA Ikemoto S., Nakamura T., Kubo M., Shimoda C.;
RT "S. pombe sporulation-specific coiled-coil protein Spo15p is localized
RL to the spindle pole body and essential for its modification.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeil C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Ruster S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Woldjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Gaffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
RA Calibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M.R., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [3]
RP SEQUENCE OF 705-871 FROM N.A.
RC STRAIN=968 H90;
RX MEDLINE=20223868; PubMed=10759889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
RL fission yeast cells by the use of a GFP-fusion genomic DNA library.";
CC Genes Cells 5:169-190(2000).
CC -!- FUNCTION: Has a role in the initiation of spore membrane
CC formation.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Spindle pole body.
CC -!- SIMILARITY: Belongs to the MPC70 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z70690; CAA94624.1; -.
CC EMBL: AB027811; BAB87115.1; -.
CC PIR: T38077; T38077.
CC GeneDB Spombe; SPAC1F3.06c; -.
CC Sporulation; Coiled coil.
CC DOMAIN 199 785 COILED COIL (POTENTIAL).
CC DOMAIN 804 1235 COILED COIL (POTENTIAL).

```

FT	DOMAIN	1320	1471	COILED COIL (POTENTIAL).
FT	DOMAIN	1481	1723	COILED COIL (POTENTIAL).
SQ	SEQUENCE	1957	AA; 222785	MM; 3F480CA06171D9DA CRC64;
Query Match 4.3%; Score 162; DB 1; Length 1957;				
Best Local Similarity 20.4%; Pred. No. 2.2;				
Matches 177; Conservative 132; Mismatches 328; Indels 232; Gaps 40;				
QY	2	VKEQNRLLN	SE-	SSQGLLGYFSDLPQ---APMVTSSTTGDLSIPSE 49
DB	129	VTQGNLNL	QVRSKLA	LEHNGILSLQLSSNKKDKNTSSVTLTSEEDVSFYQKK 188
QY	50	LENIPSEN	QYFOSAI	WS---GFIKV---KSDXYTFATSDNHNVTWVDVQEVINKAS 101
DB	189	LTMESNFS	AKOSEAY	DLRSQLLTVTEKLDKKEKYEDIKEDVSSIKASLAEQASNKS 248
QY	102	NSNKIRLE	K---GR	LIQIKIQYQENPTE---KGLDFKLWTDSQNKKEVTSNDNLQPE 155
DB	249	REQERLEK	LIVSSNK	VTSTPQTENSLRACKTLQEK-----EKCAINEDSKLEE 300
QY	156	LKQKSN	-----SR	KRSTSGAGTVPDRDNDGIPDSLEVEGYTVDVKNKR----- 200
DB	303	LXHNVANS	DAIVHKD	LIEDSLRISEFDN-----LKSERDTLSIKNEKLEKLRNTI 356
QY	201	-TFLSPW	SIHKKGL	KLYKSS---PEKWSASDPYSDPEK-----VTGRD----- 244
DB	357	GLSKDRT	SNQSEEM	VELKESNRTHISQLTDAESKLSEFGQENSKLGSIDEYQNLS 416
QY	245	-----KN	VS-----PE	ARHPLVAA-----YPIVHYDM----- 266
DB	417	SKDKMVQ	VSSQLE	EARSLAHATGKLAEBINSEDFONKKIKDFEKIEQLRACLNSSN 476
QY	267	---ENIL	SKNEQST	QNTD---SETRTISKNTS-----RHTSEVHGNAEVHASFPD 315
DB	477	ELKEKAL	DKKQDL	NLRQIQEKQKVSSESTQSSLSQRLDILNEKKKHEVYESQUNE 536
QY	316	IGGSVAG	SPNSNS	-----STVAIDHLSLAGERTWAETMGLNTADTARLNANIRYANTG 370
DB	537	LKGELO	TEISNEH	LSQSLTLAAKEAAVATNNEISEK--NSLQTL-CNA----- 585
QY	371	TAPIYNV	PTTSLV	LKGNQ---TLATIKADENQSLAPNNYPSKNLAPIALNAOKD 426
DB	586	---FQEK	LAKSVQ	LKENEQNFSSLDTSFKKLNEHQELNNHQTTKQJ-----KD 634
QY	427	FGSTPT	MYNNOF	LELEKTKQLRLDTDOVGNIAATYNFENGVRVDTGNSWSEVLPQOE 486
DB	635	TSS-----	KLQQLER	ANFQEKESTLSDENNDLRTK-----LLKLEE 672
QY	487	TTARIIF	NGKDLN	LVERRIAAVNPSPLETTKPDMLKEALKI-----AFGFNEPNGMLQ 541
DB	673	SNKSLIK	QEDVD	SLKNI-----QTLKEDLRKSEALRPFKLEAKNREVIDNL- 722
QY	542	YQKQIT	-BFDNF	QOQTSQNTKQQLAELNATNIYTVLDKIKLNAKNNILIRDKRPHYDR 600
DB	723	-KGKHET	LAQRND	LHSSLSAKNTAILSELTKSSSEVKRLTANVETLTQDSK----- 776
QY	601	NNIAVGA	DES	VVKEAHRVINSSTEGLLNLDKDIR-----KILSGYVEITE-----DT 649
DB	777	-----	ANKQST	SLVNSYQS--ISNLIHELDRDDHVMQSQNTLLESSEKLTDC 824
QY	650	EGKE---	VINDRYO	MLNISSLRQGGKTFIDFKKYNDKLPLYTISNPNYKVNIVYAKTNT 706
DB	825	ENLTQO	NMTLID	NVOKLMEKHVNQESKV-SELKEVNGKLSLDLKNLRSSLNVA-ISDNDQ 882
QY	707	IINP----	SENGDT	-----STNGIKKI 724
DB	883	ILTQLA	ELSKNY	DSLEQESQALNSGLKSL 911

RESULT 12

P115 MYCHR

ID P115 MYCHR

AC P41508;

STANDARD; PRG: 979 AA.

RESULT 12  
P115\_MYCHR  
ID\_P115  
AC P4150

```

DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE P115 protein.
OS Mycoplasma hyorhinis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OC NCBI_TaxID=2100;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=9118990; PubMed=1825306; Wise K.S.;
RA Notariicola S.M., McIntosh M.A., Wise K.S.;
RT "A Mycoplasma hyorhinis protein with sequence similarities to
RL nucleotide-binding enzymes.;"
RL Gene 97:77-85(1991).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- DOMAIN: Consists of two putative central coiled-coil regions
CC flanked by putative globular regions at the N- and C-termini.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYOPLASMA
CC P115.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M34956; AAA25423.1; -.
DR PIR: JQ0894; JC0894.
DR InterPro: IPR003439; ABC transporter.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR003405; SMC_C.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF02483; SMC_C_1.
DR Pfam: PF02463; SMC_N_1.
DR TIGRfams: TIGR00650; MG442; 2.
DR ATP-binding; Coiled coil.
DR NE_BIND: 32 39 ATP (POTENTIAL).
FT DOMAIN 159 224 COILED COIL (POTENTIAL).
FT DOMAIN 231 400 COILED COIL (POTENTIAL).
FT DOMAIN 569 821 COILED COIL (POTENTIAL).
FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 979 AA; 110566 MW; 30D51C56856280F4 CRC64;
Query Match 4.3%; Score 161.5; DB 1; Length 979;
Best Local Similarity 18.9%; Pred No 0.92;
Matches 175; Conservative 132; Mismatches 342; Indels 277; Gaps 40;
Qy 13 ESSSGGLGYFSDLNQFAPMV-----TSSTTGDLSIPSELENIPSE 56
Db 51 EQSAQLRGLNMDDVIFAGSKTKVPOEKAMKVLTPKNEDAIETKQIFITRLLKRGQGT 110
Qy 57 NOVFOF-----ATWSQFIKVKSDYEFTFATSDNHWVMDQDE--VINKASN 102
Db 111 NEYFYNDQPVRYKDKNLAVESG--ISKSSLAISQGTISEIAETPQKAVIEEAAAG 167
Qy 103 SNKILEKGRLYQTKIQYQRENTEKGLDPKLYWTDSONKKEVISDMQLPELKQKSSN 162
Db 168 TSKYKLDKEBAQKKLIR-----TNDRID-KLQGAIKELERQVNSLDK-QASKAKIYLEK 219
Qy 163 SRKKKSTAGTVPDRD-----NDGIPDSVEGYTVDVK-NKGTFTLSPMISIHKKGG 215
Db 220 SKALESVEGLIVLNDFFNEKLNLLNTSLLEVEQQRNDLELNLTQYESSISQTVH--- 275
Qy 216 LTKYKSPKSKTSADSDYDFEKTGTID--KNVSP-----EASHPLVAAYPIVHV 264
Db 276 -----FTKEVE-----SIIQETSKLNLKNALEINLQEARIEERKKLIISGEIV-V 322
Qy 265 DMENI1--LSKN-EDQSTQNTDSETRTISKNTSTSRTH----- 300
Db 323 DOKTKIEBKQVSLKIQINASKQREIELDQOLTFLNKAANSLKQENDINKIGVILE 382

```

QY 301 --SEVHGAETHASFFDGGVSAGFSNSNSTVAIDHSLSLAGERTWAE---TWGLNTA 355  
 Db 383 KSAANANILKQFENKFSKGIKTIKQNSFLFDGYIGLASELFKVESEFSLAIETV 442  
 QY 356 DPARLN-----ANIRYVNTGTAPINYVLPFTSL-----VLG 386  
 Db 443 LGAALNQIVKMTSEDLQAIIDKKNLKGKATFIPLTSIKEREVREDHLLVLKQKQGLG 502  
 QY 387 KNOTLATIKADENQ-----LSQILAPNYYPSKNLAPI-----ALN 422  
 Db 503 VAKELIEFTQNKLFGLGILNVLVVDNANRIAKILDKHYTIVSLEGLFRPGGTIT 562  
 QY 423 AQKDFSSPTIMYNYQFLEKTKQLRLDQD-----VYGNATYFNGRVRVD 472  
 Db 563 GSKGLERTSI-LNYDIKIK-EHTNLTKEADQIHLKIKQQTIVNEISTVSTIQVKIE 620  
 QY 473 TGS-----NWSEVLPQIQE--TTARILFNGKDLNVERIAAV----- 508  
 Db 621 ANSINKLNILNEELNNKLNASEIFKEQEDQESLNLSPDSEKLN-IEKQISTLIEN 679  
 QY 509 NPSDEL-----ETTKPDMTLK-----EALKIARFNGNEPNGNLQYQK 545  
 Db 680 SKKRLTNLISQKGEYKQELDAKRLKLNTOHSDSITEQNRKFLVEQNKLSHYK 739  
 QY 546 DITE-----PDFPDQOTS--QNIKNQALBNATNTYVLDKIKKANXNVLIRDKRF 596  
 Db 740 LLEASAPQYSLDLIDQARHFVDSLKKELGNNVLEATEFEENVQR-----YQBKKQ 795  
 QY 597 HYDRNNIAGADESVVKEAHREVINSSTEGLLNIDKD-----IRKILSGYIVIEDTEG 651  
 Db 796 YIEELTTAKSKEEAEISLDLKIINKTE--IVNLVNEFNVMVQKMGKGAKIHT-- 851  
 QY 652 LKEVINDRYMLNIS---SLRODGTFTDFDKKND-----KLPIYI- 689  
 Db 852 -----DKNDILNSGVEISAQPGXTIKNLRFLSGGERAKIIAISLLFALKARPIPLCIL 905  
 QY 690 -----SNPNYKNVYAVTKENT 706  
 Db 906 DEVEALDESNIWRYVEFLKIKENT 931

## RESULT 13

BXD\_CLOBO  
 AC BXD\_CLOBO STANDARD; PRT; 1276 AA.  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Botulinum neurotoxin type D precursor (EC 3.4.24.69) (BoNT/D)  
 DE (Bontoxilysin D).  
 GN BOTD.  
 OS Clostridium botulinum.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OC NCBI\_TaxID=1491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BVD/-3;  
 RX MEDLINE=91016853; PubMed=2216736;  
 RA Binz T., Kurazono H., Popoff M.R., Eklund M.W., Sakaguchi G.,  
 RA Kozaki S., Kriegstein K., Henschen A., Gill D.M., Niemann H.;  
 RT "Nucleotide sequence of the gene encoding Clostridium botulinum  
 RT neurotoxin type D.";  
 RL Nucleic Acids Res. 18:5556-5556(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CB16;  
 RX MEDLINE=93042276; PubMed=1420572;  
 RA Sunagawa H., Ohnaya T., Watanabe T., Inoue K.;  
 RA "The complete amino acid sequence of the Clostridium botulinum type D  
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding  
 RT phase d-16 phi genome.";  
 RL J. Vet. Med. Sci. 54:905-913(1992).

[3]  
 PARTIAL SEQUENCE.  
 RC STRAIN=D-SA, and D-1873;  
 RX MEDLINE=89319741; PubMed=2668193;  
 RA Morifishi K., Syuto B., Kubo S., Oguma K.;  
 RT "Molecular diversity of neurotoxins from Clostridium botulinum type D  
 RT strains";  
 RL Infect. Immun. 57:2886-2891(1989).  
 RN [4]  
 IDENTIFICATION OF SUBSTRATE.  
 RA MEDLINE=94230352; PubMed=8175689;  
 RX Yamaaki S., Baumbister A., Binz T., Blasi J., Link E., Cornille F.,  
 RA Ragues B., Fyke E.M., Suedhof T.C., Jahn R., Niemann H.;  
 RT "Cleavage of members of the synaptobrevin/VAMP family by types D and  
 RT F botulinum neurotoxins and tetanus toxin.";  
 RL J. Biol. Chem. 269:12764-12772(1994).  
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER  
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED  
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD  
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT  
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC  
 CC ENDOPEPTIDASE THAT CLEAVES THE 60-LYS-LEU-61 BOND OF  
 CC SYNAPTOSOMES-1 AND -2.  
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the  
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No  
 CC detected action on small molecule substrates.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a  
 CC heavy chain (H). The light chain has the pharmacological activity,  
 CC while the N- and C-terminal of the heavy chain mediate channel  
 CC formation and toxin binding, respectively.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of  
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.  
 CC -1- MISCELLANEOUS: BOTULINUM TYPE D NEUROTOXIN IS SYNTHESIZED BY D  
 CC STRAIN OF CLOSTRIDIUM BOTULINUM WHICH CARRY THE APPROPRIATE  
 CC BACTERIOPHAGE.  
 CC -1- SIMILARITY: Belongs to peptidase family M27.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X54254; CAA38175.1; -;  
 CC EMBL; S49407; AAB24244.1; -;  
 CC PIR; S11455; S11455.  
 CC HSP; P10845; 3BTA.  
 CC MEROPS; M27.002;  
 CC InterPro; IPR008985; ConA like lec.gl.  
 CC InterPro; IPR002160; Kunitz legume.  
 CC InterPro; IPR006025; Pept M Zn BS.  
 CC InterPro; IPR000395; Peptidase M27.  
 CC Pfam; PF01742; Peptidase M27; 1.  
 CC PRINTS; PR00760; BONTOTOXILYSIN.  
 CC ProDom; PD001963; Bontoxilysin; 1.  
 CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.  
 CC CHAIN 1 442 BOTULINUM NEUROTOXIN D, LIGHT-CHAIN.  
 CC METAL 443 1276 BOTULINUM NEUROTOXIN D, HEAVY-CHAIN.  
 CC METAL 229 229 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC METAL 230 230 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).  
 CC DISULFID 437 450 INTERCHAIN (PROBABLE).  
 CC VARIANT 15 16 ND -> PV (IN STRAIN D-1873).  
 CC VARIANT 17 18 ND -> LQ (IN STRAIN D-1873).  
 CC VARIANT 452 452 K -> Q (IN STRAIN D-SA).  
 CC VARIANT 457 457 R -> T (IN STRAIN D-SA).  
 CC VARIANT 457 457 R -> F (IN STRAIN D-1873).  
 CC VARIANT 462 462 A -> D (IN STRAIN D-1873).





```

Db 2308 YRNISETKLQMEHSTDFKPMELHKGNET-----NNKSLEKEKLLKSVN--DHMH 2359
Qy 516 TTKPDWTLKALKI-----AFGFPNPGNLQYQKQITEPDPNF--DQOTSQNKQLAEL 569
Db 2360 SMEAEM-IRKGLKYTPESVQNNIYSVIEAEVKTLEEIDRDYGDNYQVIEEHKKQFSIL 2418
Qy 570 -NATNIYTVLDKIKLNKAKMILIRDKRFHY---DRNNIAGADESVVKEAHREVINSN- 624
Db 2419 IDRTN--ALWDDHIFKFN-----NYNLEVNTETIHRVNDVIEKITNKLQVAKTE 2468
Qy 625 -EGLLNI---DKDIRKILSGYVIEIDTEGLKE---VINDRYDM-----LNISSLR 669
Db 2469 YEILENIKQNDMLQNFILFKVYSIIEYFENVKVKESILNDLYEQERLLKIGEHLEIK 2528
Qy 670 QDGKTFIDFKYNDKPLIYSN-----PNYKNVYATVKTENTLINS-----ENGDS 717
Db 2529 RVNTEIUSSEIIOQWEMSKNLLKKSQMMNY-TSIYELERANEINDAKOIKDDDTI 2587
Qy 718 TNGIKLILFSKKG 731
Db 2588 LNSVLEAAI-QKRG 2600

RESULT 15
FAT1_SCHPO STANDARD; PRT; 1385 AA.
AC O13735; Q9JUV0;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Actin interacting protein 3 homolog.
GN FAT1 OR SPAC15A10.16 OR SPAC15E1.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Gaffeau A., Cadieu E., Dreano S., Gloux S., Jelaire V., Mottier S.,
RA Calibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
RN [2]
RP IDENTIFICATION, AND GENE NAME.
RX MEDLINE=20143585; PubMed=10679021;
RA Jin H., Amberg D.C.;
RT "The secretory pathway mediates localization of the cell polarity

```

```

RT regulator Atp3p/Bud6p.";
RL Mol. Biol. Cell 11:647-661(2000).
CC -!- FUNCTION: Involved in the organization and/or function of the
CC actin cytoskeleton.
CC -!- SIMILARITY: TO YEAST BUD6.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z97208; CAB10112.1; -
DR EMBL; AL109770; CAB524420.1; -
DR GENEDB SPOMBE; SPAC15A10.16; -
DR GO; GO:0005884; C:actin filament; ISS.
DR GO; GO:0016043; P:cell organization and biogenesis; ISS.
DR InterPro; IPR005613; AIP3.
DR Pfam; PF03915; AIP3; 1.
DR Kofam; K03915; AIP3; 1.
KW Coiled coil; Cytoskeleton.
FT DOMAIN 1009 1096 COILED COIL (POTENTIAL).
SQ SEQUENCE 1385 AA; 154325 MW; ABB3D40CC4FF7537 CRC64;

Query Match 4.2%; Score 159; DB 1; Length 1385;
Best Local Similarity 19.2%; Pred. No. 2;
Matches 147; Conservative 125; Mismatches 293; Indels 202; Gaps 33;

Qy 26 DLNCAQNVVTSSTGDL-----SIPSELENIPESENYQFQSAIWS-----GFIKYKKS 74
Db 326 DRKFSPLKTRTPSLTKSLDPGTSLKSPSLRKSPPS--SPVQKDYVSRNSLRISQANRS 383
Qy 75 DEYTFATSAADNHTVMWYDDQEVINKAS-----
Db 384 NVFFCATD---VTRSVSDHRLSSSTINDGEVAPLPQRSRTISSPNSPISATVLSST 439
Qy 105 KIRLEKRLYQIKIYOQREN-----PTEKGLDFKL-----YWTDSQNK 143
Db 440 PILLPRGRSSTLSVKNKQFNADGOSTLNSPNSIRETEYAAAPKLEADIADDEVTDATSOR 499
Qy 144 EVI-----SSDNLQLPELKQSS---NSRKRSTSGPTVPDRDNGI----PDSL 187
Db 500 ELLEEQTKAESSEDTSEISLQKLSLPQVSSTQEQIOPSSSVPEAASNEIAEKPAVT 559
Qy 188 EVEGYTVDVKNKRTFLSPWISNIHEKKGLTKYKSSPEKWTASDPSDFKVTGRIDKNV 247
Db 560 AIESIT-----ERKEEAPVISSEKIESGTSISTSTKGLANPENDSLBELERLIQON- 613
Qy 248 SPEARHPILVAAPIVHVDMENIILSKVEDOSTQWTDSETRTISKNTSTSRTHTSEVHGNA 307
Db 614 NAEQDEFSYKHKYEYSSE-----SGSEDFKSEKDTKGVVISNDDSTQVEESEDKSTP 669
Qy 308 EVHASFFDIGSVSAGFSNNSSTVAID----HSLSLAGERTWAETWGLNTADTARLNAN 363
Db 670 NTGA-----SAKLINDPSSITVSDVVPKPPASPEITEPPSSALVSATSPPTNVP 720
Qy 364 I--RYVNTGTAPIYVNLPTSLVLGKQTLATIKADENQLSQILLAPNNYPSKNLAPIAL 421
Db 721 IVPEAVHLSTA--FSTAPVSTIV-----SNISPLPTVAD--P 753
Qy 422 NAKQDFSTPTTM--NTNQFLELEKT--KQLRLTDQVYGNIAFYNGRVRVYDTGSNW 477
Db 754 NVSGSPSETPISEKPEKPVVVSQTEKALPKELGVDTKEY--FLRYNNQTRKVKVESPLSN 811
Qy 478 SEVLPTQETTARIIFNGK--DLNLVERRIAANVPSPLETTKPDMT-LXEALKIAPGN 534
Db 812 ANELGELFSNVYKISFGSDSYELNIED-----PDTKISYLLUEDSDLYKSLVSPWFK 864
Qy 535 EPNGNLQ----YQGDITDFDFNFQDQTSQNIKNOLAEINATNIYT-VLDKIKLNKAKMIL 590
Db 865 EQDANKREDPHSGEVSALQHSSAQNTLDDHVNTTHESPPSAFTTEILERLKA----- 917

```

QY 591 IRDKRPHYDRNNIAGADESVVKEAHREVINSTEGLLNIDKDIRKILSGYIV-EIEDT 649  
Db 918 -----IEQNIST-----NHTNDSAAALKSSEDSKLANNFSPDSIDH 954  
QY 650 EGLKEVINDRYDMLNISSIRQDCKTFI-----DFKK-----YNDK 684  
Db 955 KFYQVKNMQ---LELASLKQISAAFFTRIPLIKIDFKKEINAFNEK 998

Search completed: May 3, 2004, 19:36:55  
Job time : 12.2539 secs